SIAM REVIEW
Vol. 44, No. 1, pp. 3–54

Combinatorial Landscapes*

Christian M. Reidys†
Peter F. Stadler‡

Abstract. Fitness landscapes have proven to be a valuable concept in evolutionary biology, combinatorial optimization, and the physics of disordered systems. A fitness landscape is a mapping from a configuration space into the real numbers. The configuration space is equipped with some notion of adjacency, nearness, distance, or accessibility. Landscape theory has emerged as an attempt to devise suitable mathematical structures for describing the "static" properties of landscapes as well as their influence on the dynamics of adaptation. In this review we focus on the connections of landscape theory with algebraic combinatorics and random graph theory, where exact results are available.

Key words. fitness landscape, genotype phenotype map, random graphs, correlation functions, neutrality, coherent algebras, sequential dynamical systems, combinatorial optimization

AMS subject classifications. 05C90, 15-02

PII. S0036144501395952

1. Introduction.

The concept of a *fitness landscape* originated in the 1930s in theoretical biology [251, 252] as a means of visualizing evolutionary adaptation. A fitness landscape is a kind of "potential function" on which a population moves uphill due to the combined effects of mutation and selection; see Figure 1.1. Implicit in this notion is both a *fitness function* f that assigns a fitness value to every possible genotype (or organism) and the arrangement of the set of genotypes in some kind of abstract space that provides some notion of accessibility or reachability.

A significant part of modern population genetics is still based on these ideas. The basic ingredients are [26] a set X of "types" (which can be either quantitative phenotypic traits (such as weight or length of limb) or a discrete genetic structure), a fitness function f evaluating the types, and a mutation function $u(x, y)$ measuring the probability that type y is obtained from type x through a mutation event. (Strictly speaking, this implies that a measure must be defined on the set of types. Since we will mostly be concerned with the case of a finite set X with the uniform measure we shall not be concerned with this complication.) A very general framework for multilocus models is described in [101].

*Received by the editors January 5, 2001; accepted for publication (in revised form) October 1, 2001; published electronically February 1, 2002. This work was supported in part by the Austrian Fonds zur Förderung der Wissenschaftlichen Forschung project P14094-MAT and LANL LDRD grant 2000515. This paper is LA-UR 00:6033.

http://www.siam.org/journals/sirev/44-1/39595.html

†Los Alamos National Laboratory D-2, MS M997, Los Alamos, NM 87545 (reidys@lanl.gov).

‡Institut für Theoretische Chemie und Molekulare Strukturbiologie, Universität Wien, Währingerstraße 17, A-1090 Wien, Austria (studla@tbi.univie.ac.at) and The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501 (external faculty).

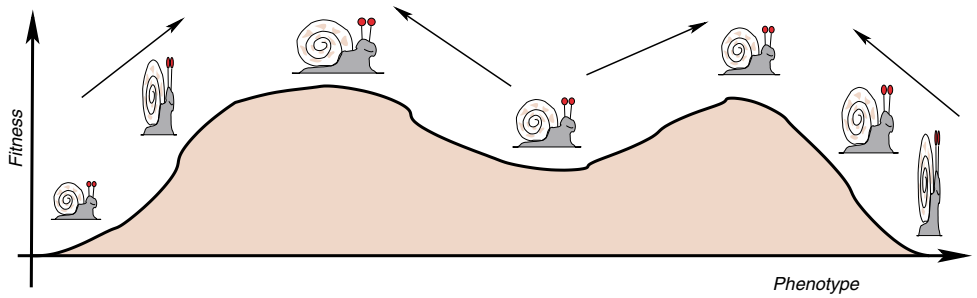


Fig. 1.1 *A fitness landscape. The fitness of this particular snail species depends on its shape. Mutations continuously produce variants that are selected if their fitness is larger than the fitness of the current "wild-type" snail. As a consequence, the shape of the snails changes over time until it reaches a maximum of the fitness landscape.*

Models of disordered systems, in particular spin glasses, naturally led to the notion of landscapes [21, 152]: Each spin configuration is assigned an energy by virtue of the Hamiltonian that specifies the model. In the simplest case so-called Ising spins are considered, which can take only two values: *up* ($\sigma = +1$) and *down* ($\sigma = -1$). The Hamiltonian of the system typically considers the interactions between neighboring spins, in the simplest case

$$f(\sigma) = \sum_{\text{neighbors } i,j} \sigma_i \sigma_j. \tag{1.1}$$

Spin configurations may change over time by flipping single spins. Of course flips that decrease the total energy $f(\sigma)$ of the system are more likely than others. We shall briefly return to this topic in section 6.

There is close conceptual similarity of the landscapes in biology and spin glass physics with the *potential energy surfaces* (PESs) of theoretical chemistry [104, 153]: As a consequence of the validity of the Born–Oppenheimer approximation, the PES provides the potential energy $U(\vec{R})$ of a molecule with n atoms as a function of its nuclear geometry $\vec{R} \in \mathbb{R}^{3n}$. Figure 1.2 show the PES of dimethoxymethane as a function of two torsion angles.

In complete analogy to this simple example, the folding of biopolymers, including proteins and nucleic acids, is determined by energy landscapes over a very large number of torsion angles. This problem was recently reviewed from a mathematical point of view in [164]. Similarly, PESs of molecular clusters can be studied [18, 48, 239]. A very recent review of the smooth energy landscapes in molecular systems is [238].

In the case of biopolymers, however, it is often useful to approximate the smooth PES by a discretized model that conceptually considers only the local minima of the PES. In the case of proteins, lattice models are prevalent [35, 47, 170] which view peptide chains as self-avoiding walks on a (usually square or cubic) lattice. The discrete standard model of RNA energy landscapes [64, 66] is briefly described later in this contribution; see section 5.2. The differences between discrete and continuous landscape models are discussed briefly in section 7.1; see also [207].

In combinatorial optimization the fitness function f is usually referred to as the *cost function* on a *search space* X [82]. The traveling salesman problem (TSP) is probably the most frequently studied combinatorial optimization problem. Hundreds

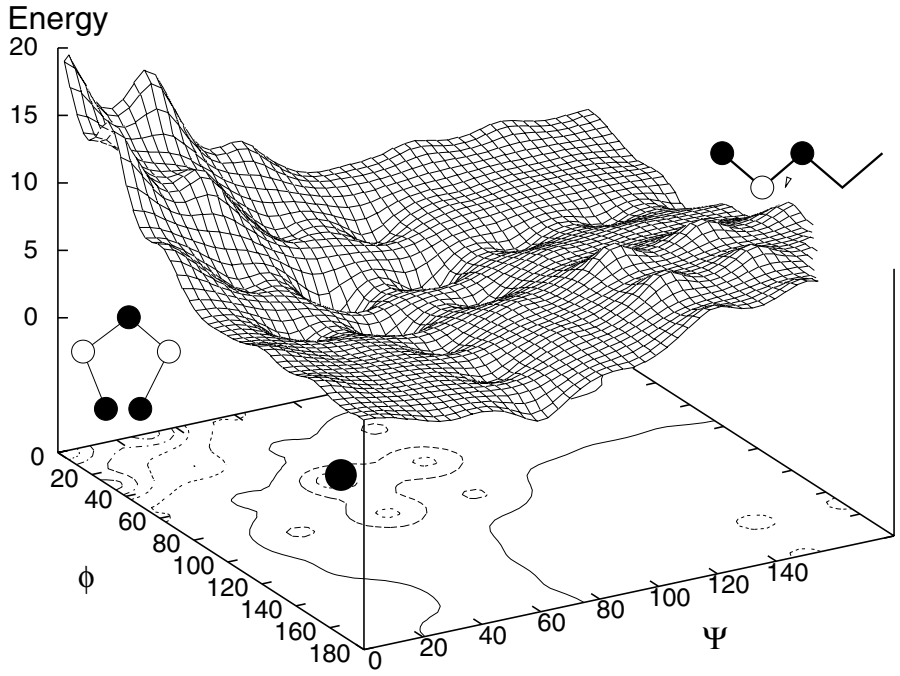


Fig. 1.2 *Potential energy surface of dimethoxymethane* $CH_3 - O - CH_2 - O - CH_3$ *as a function of the two torsion angles* ϕ *and* ψ*. The global minimum of the energy surface corresponds to preferred geometry of the molecules. Two extreme geometries, with* $\phi = \psi = 0$ *and* $\phi = \psi = 180°$*, are shown. Data provided by Alfred Karpfen.*

of publications and books have been devoted to this problem, and a large variety of solution techniques has been proposed [139]. The TSP is deceptively easy to state: *A salesman wants to visit n distinct cities and then return home. The goal is to minimize the overall traveling distance while visiting each city not more than once.* The problem is well known to belong to the class of NP-hard problems [7, 82].

The ingredients of the TSP (see Figure 1.3) are simple enough: The configurations are the $n!$ permutations of the n locations, usually called a "tour." We write $\pi = (\pi(1), \ldots, \pi(n))$ for the order in which they are visited. Given the travel distance (or cost) C_{kl} from city l to city k we can write down the cost function in the form

$$f(\pi) = \sum_{i=1}^{n-1} C_{\pi(i+1),\pi(i)} + C_{\pi(1),\pi(n)} \,, \tag{1.2}$$

where the last term describes returning to the point of origin. The TSP's apparent simplicity on the one hand and the difficulty of finding optimal solutions on the other hand have established it as a test bed for new heuristics and exact algorithms. However, the TSP is not only of theoretical value; there are many industrial applications of the TSP and its variants.

For the purpose of this contribution we are interested in the properties of the landscape of the TSP rather than in algorithms for "solving" it. A very pragmatic reason for this is that one may hope that problems with similar landscapes will lend themselves to similar solutions.

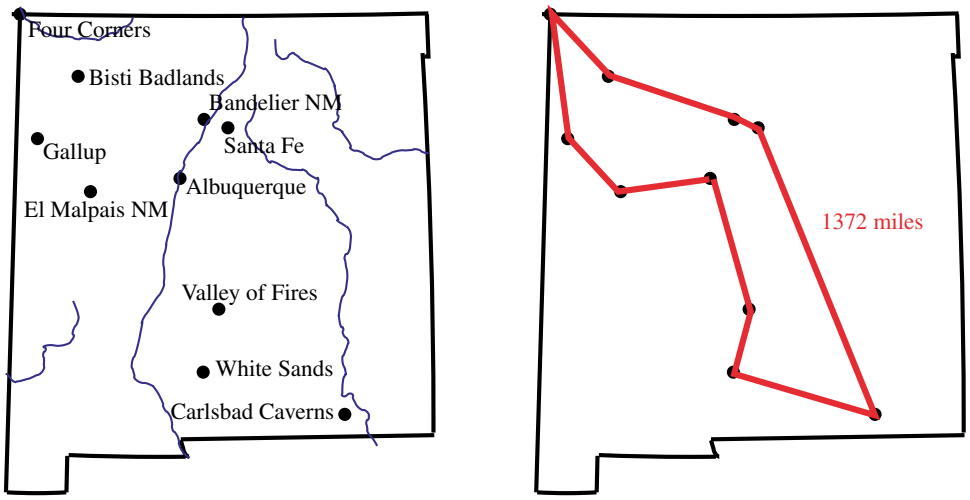


Fig. 1.3 *The "New Mexico weekend trip" is a toy example of a TSP: starting from Albuquerque airport, you rent a car in order to visit the 9 sites shown on the map. At the end, the car must be returned to Albuquerque. Of course, you want to minimize the total driving distance. An optimal solution is shown on the right-hand panel.*

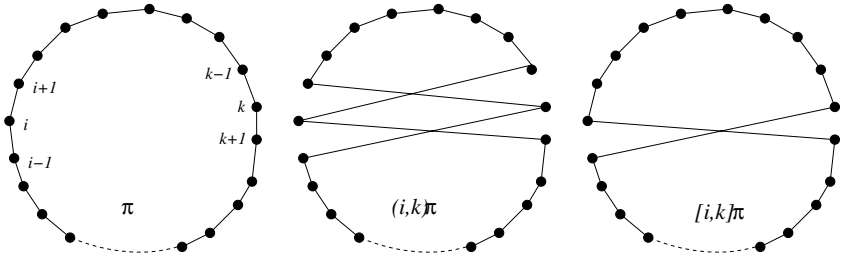


Fig. 1.4 *Transposition $(i,k)\pi$ (middle) and reversal $[i,k]\pi$ (right) of the TSP tour π (left). With a transposition two cities i and k exchange their place along the tour π; a reversal move inverts the direction of the part of the tour between i and k as well.*

A very simple approach to finding short tours is the following heuristic: Start with an arbitrary tour π_0. Then try to modify the tour, e.g., by exchanging two cities; such a *move* is called a transposition; see Figure 1.4. If the new tour π' is shorter, then retain it, otherwise try again. Of course, there is no guarantee that a heuristic procedure, such as the "adaptive walk" described above, will find the optimal solution.

Heuristic procedures operating on the landscape of a combinatorial optimization problem are by no means the only way of attacking such problems. In many cases efficient algorithms can be found that construct a globally optimal configuration from smaller parts, e.g., in dynamic programming [19] or other specialized methods [139]. Nevertheless, landscape-based heuristics such as *simulated annealing* (see section 6.1) are of substantial practical importance. The performance of these methods, both in terms of required computer resources and in terms of quality of the best configurations that are found, can be improved by tuning the move set, i.e., the way in which

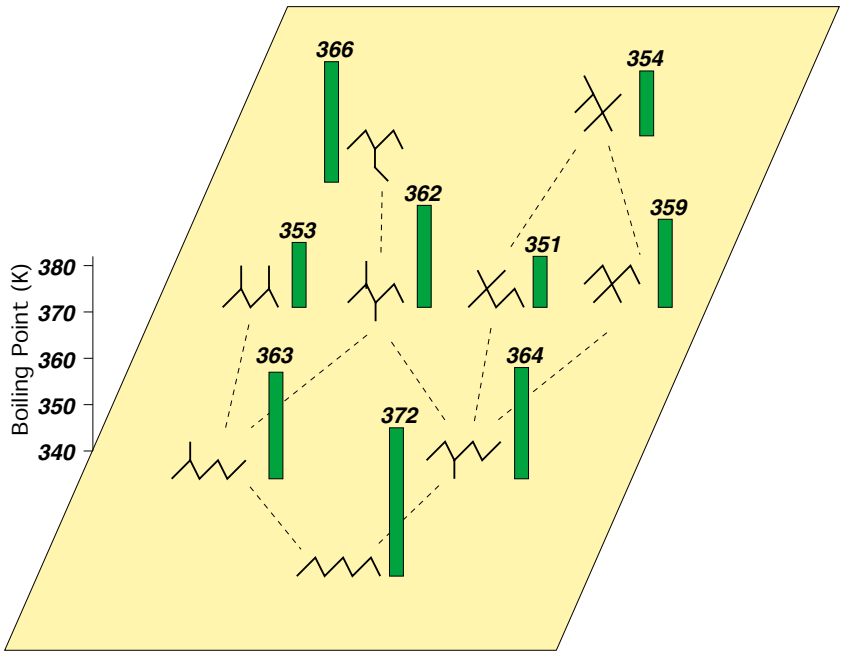


Fig. 1.5 *Boiling points of heptane isomers. The configuration space is formed by the* 9 *isomers of* C_7H_{16}*; the height of the landscape gives the boiling points of these hydrocarbons.*

configurations are modified, and by tuning the criteria for retaining a configuration. The application of evolutionary models to combinatorial optimization problems has lead to the design of so-called *evolutionary algorithms* such as genetic algorithms, evolutionary strategies, genetic programming, and memetic algorithms [69, 111, 124, 136, 151, 178].

In the case of the TSP it is an interesting and, at first glance, a rather surprising observation that simple heuristic algorithms (such as simulated annealing) work much better with "reversals" (inversion of a part of the tour; see Figure 1.4) than with transpositions as long as the matrix of travel costs is symmetric $C_{kl} = C_{lk}$ (see [143] for an analysis of asymmetric travel cost matrices). If transpositions are used, it does not matter whether the travel cost matrix is symmetric or not. An explanation for this phenomenon will be given at the end of section 3.3.

In a chemical application the configuration space typically consists of a collection of molecules (or molecular graphs, to be precise). The structure $\mathcal{X}$ arises from comparing the molecular shapes or by considering how similar the pathways of chemical synthesis are by which the molecules can be obtained in practice. Another way of arranging molecules by means of a certain partial order is described in, e.g., [130]. The "fitness" function is a particular chemical or physical property of the molecule, e.g., the boiling point or the binding constant to a cellular receptor. A very simple example is shown in Figure 1.5.

An additional complication in chemical applications is that $f(x)$ is not known for all molecules. In methods such as QSAR (quantitative structure-activity relationships) [100, 137] so-called descriptors or indices (i.e., numerical parameters) are derived from a molecular graph or via the detour of computing the detailed three-

Table 1.1 *Comparison of landscapes in different contexts.*

	Physics, Biology	Optimization	Chemistry
X	genes, genomes molecular conformations spin configurations	"configurations" e.g., TSP tours, strings	molecules
$\mathcal{X}$	mutation bond rotations spin flips **given by nature**	"move set" **designed**	similarities in synthesis or structure **designed**
f	fitness energy **known**	cost **known**	physical or chemical properties **incompletely known**
?	population structure thermodynamic properties speed of adaptation aging ...	global optimum	approximation of f then global optimum

dimensional spatial and electronic structure of the molecule. A famous example is the *Wiener index* of a graph (which is defined as the sum of all graph-theoretical distances between all atom pairs in the molecule) [97]. The vector of these indices is then related to a particular physical, chemical, or biological property of interest by means of a multivariant statistical data analysis. In other words, the first step in a QSAR application is to find an approximation f^* of f with a prescribed functional form that can be used to *predict* the property for molecules that have not yet been synthesized or investigated. In the second step the configurations (= molecules) with the optimal values of f^* are computed. These molecules are then candidates for, e.g., new drugs.

Table 1.1 summarizes some aspects of landscapes in various scientific disciplines. More recently, the language of landscapes has also been transferred to social sciences [134, 140]. For instance, *electoral landscapes* are used to describe the expected fraction of votes for a political party given the position of the voters and competing parties [134, 206]. Not surprisingly, fitness landscapes have emerged as one of the unifying themes in the literature on complex systems [38, 78, 124, 174].

Let us now turn to the main purpose of the review, the mathematical theory of fitness landscapes. In formal terms, a *landscape* is a triple $(X, \mathcal{X}, f)$ consisting of

1. a set X of configurations;
2. a notion $\mathcal{X}$ of neighborhood, nearness, distance, or accessibility on X; and
3. a fitness function $f : X \to \mathbb{R}$.

The set X together with the "structure" $\mathcal{X}$ forms the configuration space. The definition of $\mathcal{X}$ is purposefully left vague at this point and will be made precise in the following section.

Landscapes can be studied either from a "static" point of view, focusing on geometric properties such as smoothness, ruggedness, and neutrality, or from a "dynamical" point of view, focusing on the features of a dynamical system, for instance, an evolving population, that uses the landscape as its substrate. The static point of view will be the main topic of this review as it lends itself readily to a detailed mathematical analysis. Dynamical aspects are much more difficult to analyze and are mostly tackled by computer simulations.

This contribution is organized as follows: We first identify the structure of configuration spaces either as undirected, unweighted graphs or as reversible Markov chains.

This sets the stage for decomposition of landscapes in terms of particular orthonormal bases that take into account the structure of the underlying configuration space. This "spectral" approach concentrates on ruggedness, neutrality, and isotropy and is of particular relevance for combinatorial optimization problems and disordered systems. Landscapes arising in biology are based upon an underlying genotype-phenotype map which determines key features of the landscape. We discuss two paradigmatic examples of genotype-phenotype maps: RNA secondary structure folding and sequential dynamical systems. The analysis of these examples naturally leads to a random graph theory of neutrality. In section 6 we briefly review dynamical aspects in landscape theory, in particular, simulated annealing and the quasi-species model. We close our discussion with a few remarks on recent trends.

2. Configuration Spaces. In this section we briefly review the most common ways of organizing the set of configurations into a *configuration space*. Since most of this deals with standard constructions we keep it as brief as possible. There are three major structures $\mathcal{X}$ on X that are used in landscape theory:

1. Sometimes transition probabilities are specified that describe how frequently a system attempts to move from one configuration to another. This is the usual situation in a population genetics setting, where a mutation function $u(x, y)$ is given.
2. In computer science one typically specifies a "move set" or "genetic operator" that interconverts one or more configurations into a new one. Evolutionary biology takes the same point of view, except that the move sets are given by nature rather than engineered.
3. A rigorous mathematical analysis often starts with specifying a metric or a topology on X. This approach will be discussed in some detail in section 7.1.

We shall see below that move sets and transition probabilities are closely related. Eventually, we obtain an algebraic description of the configuration space in terms of a matrix that may serve as the starting point of *spectral landscape theory*.

2.1. Markov Chains. Regarding X as a set of "states" we may specify transition probabilities $\mathbf{T}_{xy}$ for moving from y to x. The Markov process with transition matrix $\mathbf{T}$ organizes the configuration space. Typically, one requires $\mathbf{T}$ to be ergodic (i.e., every state can be reached from every other state) and reversible, i.e., to satisfy

(E) $\mathbf{T}$ is irreducible, or, equivalently,
there is a unique stationary distribution p on X such that $\mathbf{T}p = p$. Furthermore $p(x) > 0$ for all $x \in X$.

(R) $\mathbf{T}_{xy}p(y) = \mathbf{T}_{yx}p(x)$. This condition is also known as "detailed balance."

In other words, $\mathbf{T}$ is self-adjoined w.r.t. the scalar product

$$\langle f, g\rangle_p = \sum_x p(x) f(x) g(x)^*, \tag{2.1}$$

where $*$ denotes complex conjugation. The most important examples in landscape theory are transition matrices associated with mutation and recombination, which are discussed in the following section in more detail.

A most useful observation is that the matrix $\mathbf{S}$ defined by

$$\mathbf{S}_{xy} = p(x)^{-1/2}\mathbf{T}_{xy}p(y)^{1/2} \tag{2.2}$$

is symmetric and similar to $\mathbf{T}$; see, e.g., [168]. Hence given a nonsymmetric transition matrix $\mathbf{T}$ and a landscape f we may transform the model to new coordinates with

the symmetric operator $\mathbf{S}$ and the transformed landscape

$$f^{\sigma}(x) = p(x)^{-1/2} f(x). \tag{2.3}$$

This allows for the application of much of the spectral landscape theory in the non-symmetric case as well.

2.2. Move Sets. In its most abstract form a *move set* assigns to a k-tuple $(x_1, \ldots, x_k) \in X^k$, which we refer to as "parents," a list $N(x_1, \ldots, x_k) \subseteq X$, which we call "children." In the following we will restrict our attention to the two most commonly used move set types, which are mutation and recombination.

2.2.1. Mutation. A mutation operator simply assigns a set $N(x)$ of "accessible neighbors" or "elementary mutants" to each configuration x; see Figure 2.1. This allows us to interpret X as a (possibly directed) graph with vertex set X and $N(x)$ the (out)neighbors of $x \in X$. Most commonly the move sets are constructed such that $y \in N(x)$ if and only if $x \in N(y)$, i.e., the graph is undirected. A graph is faithfully represented by its adjacency matrix $\mathbf{A}$, which has the entries $\mathbf{A}_{xy} = 1$ if $x \in N(y)$ and $\mathbf{A}_{xy} = 0$ otherwise. Obviously, $\mathbf{A}$ is symmetric if and only if the graph is undirected. With each (directed or undirected) graph there is an associated Markov process on its vertex set [145] defined by the transition matrix

$$\mathbf{T} = \mathbf{A}\mathbf{D}^{-1}, \tag{2.4}$$

where $\mathbf{D}$ is the so-called degree matrix, which is diagonal, and $\mathbf{D}_{xx} = |N(x)|$ is the number of neighbors of x. This Markov process describes a *random walk* on X which has been suggested as a means to sample information about a landscape by Weinberger [242, 243]. We remark that in the case of undirected and symmetric directed graphs the stationary distribution is given by

$$p(x) = \frac{\mathbf{D}_{xx}}{2E}, \tag{2.5}$$

where E is the total number of undirected edges.

2.2.2. Recombination and Crossover. The most immediate consequence of the fact that recombination acts on two arguments is that the recombination-induced configuration space cannot be represented as a simple graph with the set of genotypes representing the set of vertices. This leaves two alternatives: One can change the nature of the vertex set and have pairs of types as vertices. Then one obtains again a (di)graph, since each elementary recombination event creates up to two different strings. This approach was pioneered by Culberson [39] and Jones [119]. The alternative is to represent individual genotypes by vertices and to make the edges more complex. In Gitchoff and Wagner [86] it was shown that recombination spaces can be represented as hypergraphs (which consist of a vertex set X and a collection $\mathcal{E}$ of (not necessarily) distinct subsets of X called (hyper)edges), where the hyperedges are the sets of all recombinants that can arise from the recombination of two types. With this approach it was easy to show that string recombination spaces and point mutation spaces are homomorphic. Hypergraphs are still not completely satisfactory, since they do not indicate which pair of types produces which set of recombinants, i.e., which hyperedge arises from which mating. This led us to invent *P-structures* $P : X \times X \to 2^X$, which are mappings of pairs of types to the hyperedges of the hypergraph [221, 237].

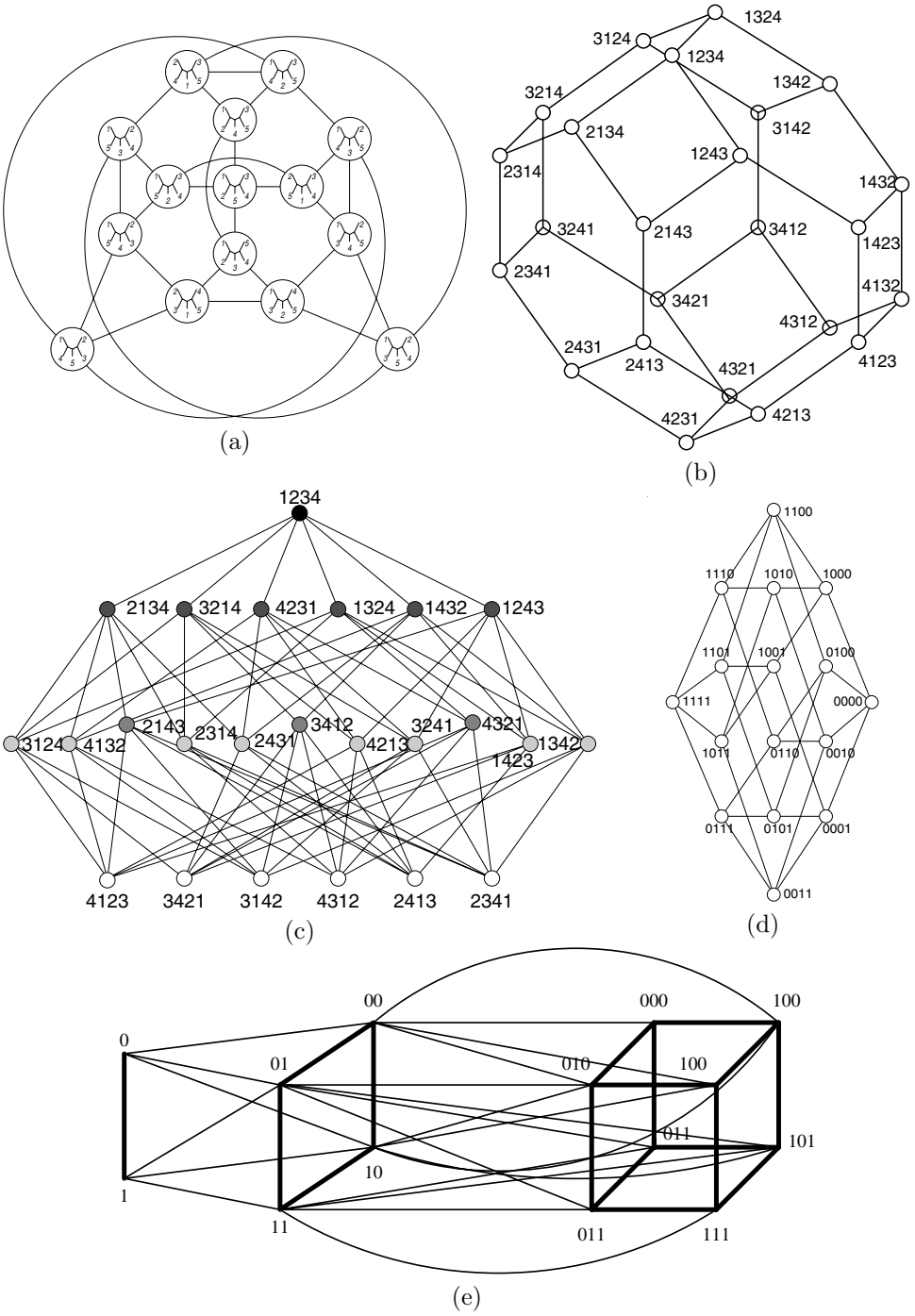


Fig. 2.1 *Some examples of configuration spaces.* (a) *Phylogenetic trees with five taxa related by so-called nearest neighbor interchanges.* (b) *The permutohedron graph arises from the permutations of* (1,2,3,4) *that differ by transposing two* adjacent *entries.* (c) *The same permutations, but this time all transpositions lead to neighbors.* (d) *The sequence space (Hamming graph) of binary sequences of length* 4. (e) *The sequence space of all binary sequences with length at most* 3. *Moves are mutations* $0 \leftrightarrow 1$ *as well as insertions and deletions of single* 1*'s and* 0*'s.*

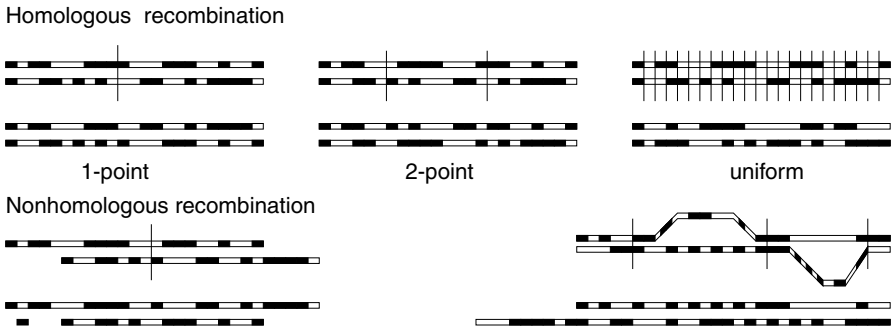


Fig. 2.2 *Recombination mechanisms. Homologous recombination works on aligned sequences of equal length. The number of crossover points may vary. In the extreme case of uniform recombination each locus is transmitted independently to the offspring genes. In contrast, nonhomologous recombination works on imperfectly aligned sequences.*

We focus here on homologous recombination on a genome consisting of n loci; see Figure 2.2. For each locus k, there are α_k alleles. The set of all the $\prod_k \alpha_k$ possible genotypes will be denoted by X. For each locus k, we label the alleles using a letter from the alphabet $\mathcal{A}_k = \{0, \ldots, \alpha_k - 1\}$. Thus $X = \prod_k \mathcal{A}_k$. A particular genotype (or sequence) $x \in X$ can be regarded as a vector with components $x_k \in \mathcal{A}_k$. A particular crossover operator is determined by the list χ of loci that the child inherits from the first parent. Thus the loci in $\overline{\chi} = \{1, \ldots, n\} \setminus \chi$ come from the second parent. More formally, given χ, the offspring $x = \chi(y, z)$ of the two parents y and z has the componentwise representation

$$x_k = \begin{cases} y_k & \text{if} \quad k \in \chi, \\ z_k & \text{if} \quad k \in \overline{\chi}. \end{cases} \tag{2.6}$$

It will be convenient in the following to express (2.6) by means of an "incidence operator":

$$\mathbf{H}^{\chi}_{x,(y,z)} = \begin{cases} 2 & \text{if} \quad x = y = z, \\ 1 & \text{if} \quad y \neq z \\ 0 & \text{otherwise}\,. \end{cases} \qquad \text{and} \quad x = \chi(y, z) \tag{2.7}$$

Here we restrict ourselves to recombination on strings. Crossover operators for permutations, such as traveling salesman tours, are reviewed, for instance, in [138].

A *recombination operator* in the sense of most of the genetic algorithm (GA) literature is then a family $\mathcal{F}$ of crossover operators that act on $X \times X$ with probability $\pi(\chi)$. The incidence "matrix" associated with a recombination operator is simply

$$\mathbf{H}^{\mathcal{F}} = \sum_{\chi \in \mathcal{F}} \mathbf{H}^{\chi}. \tag{2.8}$$

The two most important recombination operators are the following:

- [∞] Uniform recombination contains all 2^n possible crossover operators. In this case it is natural to include the identity $\imath$.
- [1] 1-point recombination contains all crossover operators χ for which the characteristic set is of the form $\chi = \{1, \ldots, k\}$.

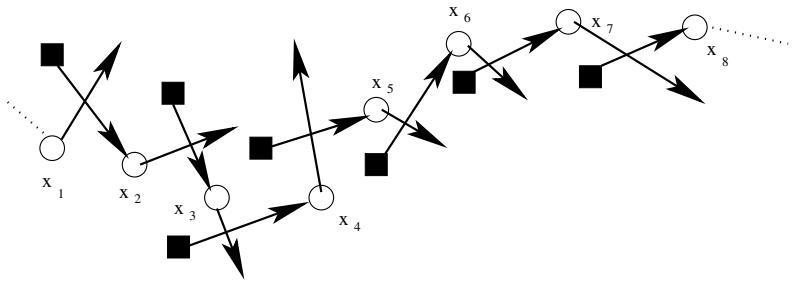


Fig. 2.3 *Crossover walk. The "father" (square) is mated with a randomly chosen "mother" (circle). One of the offsprings is the "son" (next square), which becomes the "father" of the next mating.*

Homologous recombination (of strings) under very general conditions leads to very regular configuration spaces. In particular, one can show that the automorphism group of $\mathbf{H}^{\mathcal{F}}$ is generously transitive [221]. This picture, however, changes radically if unequal crossover is considered, where the number of genes on a chromosome can change [203].

A *crossover walk* [112, 113] (see Figure 2.3) on X is the Markov process based on the following rule: The "father" y is mated with a randomly chosen "mother" z. The offspring is "son" x, which becomes the "father" of the next mating. We regard the sequence of "fathers" as a random walk on X. It is straightforward [219] to derive the transition matrix of this Markov process for homologous recombination from the incidence "matrix" $\mathbf{H}^{\mathcal{F}}$. One obtains

$$\mathbf{S}_{xy}^{\mathcal{F},\vec{\wp}} = \sum_{\chi\in\mathcal{F}} \pi(\chi)\frac{1}{2}\sum_{z\in X} \mathbf{H}^{\chi}_{x,(y,z)}\wp(z), \tag{2.9}$$

where $\wp(z)$ denotes the frequency distribution of the genotypes in the equilibrium population.

3. Decompositions of Landscapes. Regarding $f : X \to \mathbb{R}$ as a vector in the $|X|$-dimensional Euclidean vector space $\mathbb{R}^{|X|}$ immediately poses the question of whether there are more convenient bases than the standard basis $\{\delta_x\}$, with $\delta_x(y) = 1$ if $y = x$, and 0 otherwise. This is the starting point of what one might call *algebraic landscape theory* or *spectral landscape theory.*

3.1. Fourier Transform of Landscapes. A suitable basis naturally encapsulates information about the regularities of the configuration space. Hence one of the symmetric $X \times X$ matrices introduced in the previous section is the most common starting point.

3.1.1. Discrete Laplace Operators. From the algebraic point of view it appears to be more natural to start with a discrete *Laplace operator*

$$-\mathbf{\Delta} = \mathbf{D_S} - \mathbf{S} \qquad \text{with } (\mathbf{D_S})_{xx} = \sum_{y\in X} \mathbf{S}_{xy} \tag{3.1}$$

since it has a number of desirable mathematical properties:

(1) $-\mathbf{\Delta}$ is symmetric and has nonpositive off-diagonal entries.
(2) $-\mathbf{\Delta}$ has 0 as an eigenvalue with eigenvector $1 = (1, \dots, 1)$. The eigenvalue 0 is unique if and only if the graph associated with the off-diagonal entries is irreducible.
(3) $-\mathbf{\Delta}$ is nonnegative definite.

The graph Laplacian $-\mathbf{\Delta} = \mathbf{D} - \mathbf{A}$ arises naturally as the discretization of the Laplacian differential operator, for instance, in finite element computations. For recent surveys on graph Laplacians, see [36, 150, 157, 158].

3.1.2. Coherent Algebras. An alternative, or maybe an even more appealing starting point, is to consider the coherent algebra associated with the configuration space graph or transition operator [131, 220].

A set of complex matrices that is closed under (i) scalar multiplication with complex numbers, (ii) componentwise addition, (iii) ordinary matrix multiplication, (iv) componentwise multiplication, and (v) transposition is called a *coherent algebra* or *cellular algebra*. Equivalently, a matrix algebra $\mathfrak{W} \subseteq \mathbb{C}^{|X|\times|X|}$ is coherent if and only if it satisfies the following axioms:

(i) As a linear space over $\mathbb{C}$, $\mathfrak{W}$ has a basis of $\{\mathbf{R}^{(1)}, \dots, \mathbf{R}^{(r)}\}$ of 0-1 matrices.
(ii) $\sum_{j=1}^{n} \mathbf{R}^{(j)} = \mathbf{J}$, the all-1 matrix.
(iii) For every $i \in \{1, \dots, r\}$ there is an i' such that $\mathbf{R}^{(i)\mathsf{T}} = \mathbf{R}^{(i')}$.
(iv) $\mathbf{I} \in \mathfrak{W}$.

Axiom (ii) above implies that the relations associated with the basis matrices $\mathbf{R}^{(j)}$ form a partition of $X \times X$. Such partitions are known as *coherent configurations* [106, 107, 108]. These are highly regular structures that combinatorially behave almost like groups. The main difference is that they do not require an algebraic operation (multiplication) to be defined on X. Indeed, in the context of fitness landscapes a group operation on X appears rather contrived: What would we mean, e.g., by the "multiplication" or "addition" of two genes?

For each collection $\mathcal{M} = \{\mathbf{M}_1, \dots, \mathbf{M}_k\}$ of $|X| \times |X|$ matrices there is a smallest coherent algebra $\langle\langle \mathcal{M} \rangle\rangle$, which is defined as the intersection of all coherent algebras that contain $\{\mathbf{M}_1, \dots, \mathbf{M}_k\}$. Since the centralizer algebra is coherent we have

$$\langle\langle \mathcal{M} \rangle\rangle \subseteq \mathfrak{V}_{\mathbb{C}}(\mathsf{Aut}[\mathcal{M}], X). \tag{3.2}$$

Equality holds if and only if there is a permutation group that has $\langle\langle \mathcal{M} \rangle\rangle$ as its centralizer algebra [132]. The coherent algebra $\langle\langle \mathcal{M} \rangle\rangle$ can therefore be regarded as a "combinatorial approximation" of the centralizer algebra [56, 131]. This is of particular importance in the graph case: given the adjacency matrix $\mathbf{A}$ of Γ, there is a polynomial time algorithm that determines the coherent algebra $\mathfrak{W}(\Gamma) = \langle\langle \mathbf{A} \rangle\rangle$; see [9, 10, 245]. It is straightforward to check that the degree matrix $\mathbf{D}$, and hence also the transition operator $\mathbf{T} = \mathbf{A}\mathbf{D}^{-1}$ and the associated Laplace operators, are contained in $\langle\langle \mathbf{A} \rangle\rangle$.

Let $\mathcal{R} = \{\mathbf{R}^{(1)}, \dots, \mathbf{R}^{(r)}\}$ be the standard basis of a coherent algebra $\mathfrak{W}$. We have $\mathbf{R}^{(\mu)}\mathbf{R}^{(\nu)} = \sum_{\kappa} p^{\kappa}_{\mu,\nu} \mathbf{R}^{(\kappa)}$, where the *intersection numbers*

$$p^{\kappa}_{\mu,\nu} = \left|\left\{ z \in X \,\middle|\, (x,z) \in \mu \text{ and } (z,y) \in \nu \right\}\right| \in \mathbb{N}_0 \tag{3.3}$$

are the same for all pairs $(x, y) \in \kappa$. The $r \times r$ matrices $\hat{\mathbf{R}}^{\kappa}$ with entries $\hat{\mathbf{R}}^{(\nu)}_{\mu,\kappa} = p^{\kappa}_{\mu,\nu}$ generate a matrix algebra $\hat{\mathfrak{W}}$ that is isomorphic to $\mathfrak{W}$ [106]. This observation makes coherent algebras appealing objects for our purposes because $\hat{\mathfrak{W}}$ is small enough in

many cases to allow for explicit computations of eigenvalues and eigenvectors [220]. This is of particular interest in the case of association schemes (symmetric coherent algebras), in which case all members of $\widehat{\mathfrak{W}}$ share a common orthonormal basis of eigenvectors [25]. We remark that, in the case of Hamming graphs, the eigenvectors are the well-known Walsh functions [62].

3.1.3. Graph-Theoretical Fourier Decomposition. In the following we will let $\{\varphi_k | X \to \mathbb{C}\}$ be an orthonormal basis of eigenfunctions of the desired operator taken from $\mathfrak{W}$. It is then natural to expand the fitness function f in terms of this basis:

$$f(x) = \sum_k a_k \varphi_k(x). \tag{3.4}$$

We shall use the following convention: (i) The index 0 is reserved for the "ground state." If the basis is derived from a Laplacian, for instance, then φ_0 is constant, the associated eigenvalue is zero, and

$$a_0 = \sum_x \varphi_0(x) f(x) = |X|^{-1} \sum_{x \in X} f(x). \tag{3.5}$$

(ii) The distinct eigenvalues of $-\mathbf{\Delta}$ will be denoted by Λ_p and in the Markov chain case we write λ_p. It will be convenient to define the index sets $J_p = \{k | -\mathbf{\Delta}\varphi_k = \Lambda_p \varphi_k\}$ that collect all eigenfunctions belonging to the same (Laplacian) eigenvalue. (iii) We write $\tilde{f}(x) = f(x) - a_0$. This is the "nonflat" part of the fitness function.

3.1.4. Group-Theoretical Fourier Transformation. Let G be a finite group and let S be a symmetric set of generators of G, i.e., $\langle S \rangle = \mathsf{G}$, $S = S^{-1}$, and $\imath \notin S$, where $\imath$ is the identity of G. A graph $\Gamma(\mathsf{G}, S)$ with vertex set G and edges $\{s, t\}$ if and only if $t^{-1}s \in S$ is called a *Cayley graph*. Cayley graphs are vertex transitive and hence regular.

DEFINITION 3.1. *A Cayley graph* $\Gamma(\mathsf{G}, S)$ *is called* quasi-Abelian *if* S *is the union of some conjugacy classes of* G.

Clearly, a Cayley graph on a commutative group is quasi-Abelian, since each group element forms its own conjugacy class in this case. Some interesting properties of quasi-Abelian Cayley graphs are discussed in [240, 253]. Below we shall see that certain algebraic properties of Cayley graphs with Abelian groups generalize to quasi-Abelian Cayley graphs.

DEFINITION 3.2. *For any function* $f : \mathsf{G} \to \mathbb{C}$ *and any matrix representation* $\varrho = \{\rho(s)\}_{s \in \mathsf{G}}$ *of* G *we call the matrix sum*

$$\widehat{f}(\varrho) = \sum_{x \in \mathsf{G}} f(x) \rho(x) \tag{3.6}$$

the (group theoretic) Fourier transform *of* f *at* ϱ.

In the case of Cayley graphs we therefore have to distinguish between the "Fourier series expansion" w.r.t. the graph $\Gamma(\mathsf{G}, S)$, (3.4), and the representation theoretical Fourier transformation on the group G itself. It will not come as a surprise that there is an intimate connection between the two.

THEOREM 3.3 (see [192]). *Let* $\Gamma(\mathsf{G}, S)$ *be a quasi-Abelian Cayley graph with a finite group* G.

Table 3.1 *Elementary landscapes.*

Problem	Graph	D	λ	Order	Reference
p-spin glass	$\mathcal{Q}_2^n$	n	$2p$	p	definition
NAES	$\mathcal{Q}_2^n$	n	4	2	[94]
Weight partitioning	$\mathcal{Q}_2^n$	n	4	2	[94, 211]
GBP (constrained)	$\mathcal{Q}_2^n$	n	4	2	[6]
Max cut	$\mathcal{Q}_2^n$	n	4	2	[6]
Graph α-coloring	$\mathcal{Q}_2^\alpha$	$(\alpha-1)n$	2α	2	[211]
XY-spin glass	$\mathcal{Q}_\alpha^n$	$(\alpha-1)n$	2α	2	[81]
for $\alpha > 2$:	$\mathcal{C}_\alpha^n$	2	$8\sin^2(\pi/\alpha)$	2	[81]
Linear assignment	$\Gamma(\mathcal{S}_n, \mathcal{T})$	n		1	[192]
TSP symmetric	$\Gamma(\mathcal{S}_n, \mathcal{T})$	$n(n-1)/2$	$2(n-1)$	2	[37, 94]
	$\Gamma(\mathcal{S}_n, \mathcal{J})$	$n(n-1)/2$	n	2	[37, 94]
	$\Gamma(\mathcal{A}_n, \mathcal{C}_3)$	$n(n-1)(n-2)/6$	$(n-1)(n-2)$	?	[37]
antisymmetric	$\Gamma(\mathcal{S}_n, \mathcal{T})$	$n(n-1)/2$	$2n$	3	[211]
	$\Gamma(\mathcal{S}_n, \mathcal{J})$	$n(n-1)/2$	$n(n+1)/2$	$\mathcal{O}(n)$	[211]
Graph matching	$\Gamma(\mathcal{S}_n, \mathcal{T})$	$n(n-1)/2$	$2(n-1)$	2	[211]
Graph bipartitioning	$J(n, n/2)$	$n^2/4$	$2(n-1)$	2	[94, 215, 217]

(1) *The function* $\varepsilon_{ij}^k : \mathsf{G} \to \mathbb{C}$ *defined as*

$$\varepsilon_{ij}^k(u) = \sqrt{\frac{d_k}{|\mathsf{G}|}}\rho_{ij}^k(u^{-1}) \tag{3.7}$$

is an eigenvector of $\mathbf{A}(\Gamma)$ *with eigenvalue* $\Lambda_k = \frac{1}{d_k}\sum_{s\in S}\chi_k(s)$, *where* $\chi_k(s) = \mathrm{Tr}\rho^k(s)$ *is the character of* ϱ^k *at* s; *its dimension is* d_k.

(2) *All quasi-Abelian Cayley graphs on* G *have a common basis of eigenvectors, and hence their adjacency matrices commute.*

(3) *A function* $f : \mathsf{G} \to \mathbb{R}$ *can be expanded in the form*

$$f(s) = \sum_{ijk} a_{ij}^k \varepsilon_{ij}^k(s) \qquad \textit{with} \qquad a_{ij}^k = \sqrt{\frac{d_k}{|\mathsf{G}|}}\,\widehat{f}_{ji}(\rho^k). \tag{3.8}$$

Fast Fourier transform algorithms are known for a variety of finite groups. For a recent overview, see, e.g., [149, 191].

3.1.5. Elementary Landscapes. Grover and others [37, 94, 211] observed that $\tilde{f}$ is in many cases an eigenfunction of the graph Laplacian $-\mathbf{\Delta}$ (see Table 3.1 for a list of examples). We say that f is *elementary* w.r.t. $-\mathbf{\Delta}$ if $\tilde{f}$ is an eigenfunction of $-\mathbf{\Delta}$ with an eigenvalue $\lambda_p > 0$. In [212] this notion is extended to calling f elementary w.r.t. a random walk transition operator iff $\mathbf{S}\tilde{f} = \lambda_p\tilde{f}$ with an eigenvalue $\lambda_p < 1$.

If f is elementary, then $\tilde{f}$ satisfies the conditions of Courant's nodal domain theorem; see section 3.1.6. Elementary landscapes can thus be expected to have few nodal domains if they belong to a small Laplacian eigenvalue (or to an eigenvalue of a Markov transition matrix close to 1), while landscapes that are far away from the ground state will in general have many nodal domains. Such landscapes will appear "rugged." Grover [94] showed that

$$f(\hat{x}_{\min}) \le a_0 \le f(\hat{x}_{\max}), \tag{3.9}$$

where $\hat{x}_{\min}$ and $\hat{x}_{\max}$ are arbitrary local minima and maxima, respectively. This *maximum principle* shows that elementary landscapes are well behaved: There are no local optima with worse than average fitness. We shall return to local optima as a measure of ruggedness in section 4.2.1.

3.1.6. The Nodal Domain Theorem. Discrete Schrödinger operators are defined as

$$\mathbf{H}f(x) = \sum_{y\sim x} b(x,y)\,[f(x) - f(y)] + v(x)f(x), \tag{3.10}$$

where $b(x,y) = b(y,x) > 0$ if and only if $\{x,y\}$ is an edge of the graph Γ; $v(x)$ is an arbitrary "potential function." Of course, the graph Laplacian $-\mathbf{\Delta}$ is a special case.

A well-known feature of Schrödinger operators on Riemannian manifolds M is that the nodal domains of their eigenfunctions f, that is, the connected components of $M \setminus f^{-1}(0)$, are severely constrained. In order to formulate *Courant's theorem* for graphs, we define for any function $f : X \to \mathbb{R}$ on Γ: $\operatorname{supp}_+(f) = \{x \in X | f(x) > 0\}$, $\operatorname{supp}_-(f) = \{x \in X | f(x) < 0\}$, $\operatorname{zero}(f) = \{x \in X | f(x) = 0\}$, $\operatorname{supp}^0_+(f) = \operatorname{supp}_+(f) \cup \operatorname{zero}(f)$, and $\operatorname{supp}^0_-(f) = \operatorname{supp}_-(f) \cup \operatorname{zero}(f)$. A (strong) *nodal domain* of f is a maximal connected component of either $\operatorname{supp}_+(f)$ or $\operatorname{supp}_-(f)$. A *weak nodal domain* is a maximal connected component of $\operatorname{supp}_+(f) \cup \operatorname{zero}(f)$ or $\operatorname{supp}_-(f) \cup \operatorname{zero}(f)$, respectively.

Let $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_{|X|}$ be the eigenvalues of a Schrödinger operator on Γ with corresponding eigenvectors φ_i. Define $M(i) = \max\{k | \lambda_k = \lambda_i\}$ and $m(i) = \min\{k : \lambda_k = \lambda_i\}$. Hence, $m(i) \le i \le M(i)$, $M(i) = m(i) + \operatorname{mult}(\lambda_i) - 1$, and $m(i) = M(i) = i$ if and only if λ_i is a simple eigenvalue of $\mathbf{H}$. With this notation we have the following theorem.

THEOREM 3.4 (nodal domain theorem). *Let f be an eigenvector of $\mathbf{H}$ with eigenvalue λ_i. Then*

(i) *there are at most $M(i)$ (strong) nodal domains of f, and*
(ii) *there are at most $m(i)$ weak nodal domains of f.*

Various discrete versions of the nodal domain theorem have been discussed in the literature [44, 49, 80, 230], however, sometimes with ambiguous statements and incomplete or flawed proofs. An elementary proof can be found in [41]. The interest in nodal domain properties of graph eigenvectors goes back to Fiedler [60], who showed that the number of components of $\operatorname{supp}^0_+(f)$ is at most $M(i)$. Some closely related results on the component structure of $\operatorname{supp}_+(f) \cup \operatorname{supp}_-(f)$ can be found in [176].

3.2. Random Landscapes. In many cases, for instance, in applications to spin glasses, the definition of the landscape contains a number of random parameters. We therefore define landscapes here as elements of an appropriate probability space.

DEFINITION 3.5. *Let X be a finite set and let W be a predicate of landscapes $f : X \to \mathbb{R}$. A* random W-landscape *$\mathcal{F}$ over X is the probability space*

$$\Omega = (\{f : X \to \mathbb{R} \mid f \ \textit{has property } W\}, \mathcal{A}, \mu), \tag{3.11}$$

where $\mathcal{A}$ is a σ-field and $\mu : \mathcal{A} \to [0,1]$ is a measure.

Let $\xi : \Omega \to \mathbb{R}$ be a Ω-random variable; we denote expectation value and variance of ξ by $\mathbb{E}[\xi]$ and $\mathbb{V}[\xi]$, respectively. In particular we will consider the family

$$\forall\, x \in X; \quad \operatorname{eval}_x : \Omega \to \mathbb{R}, \quad \operatorname{eval}_x(f) = f(x). \tag{3.12}$$

By abuse of notation we shall write $\mathbb{E}[f(x)]$ for $\mathbb{E}[\operatorname{eval}_x]$, the expectation of f evaluated at $x \in X$. The *covariance matrix* of the random landscape Ω is given by

$$\mathbf{C}_{xy} = \operatorname{Cov}[\operatorname{eval}_x, \operatorname{eval}_y] = \mathbb{E}[f(x)f(y)] - \mathbb{E}[f(x)]\mathbb{E}[f(y)]. \tag{3.13}$$

Table 3.2 *Examples of additive random landscapes. The component landscapes ϑ_I and the index set M,* (3.16)*, are listed together with information on whether the models are uniform (U), strictly uniform (S), or pseudo-isotropic (P). Properties that are implied by stronger ones are shown as ∘. Summarized from the discussion in* [188].

Model	Component landscapes and index set	U	S	P
Ising spin glass	$\vartheta_I(x) = \prod_{k\in I} x_k \qquad I \subseteq \{1,\dots,n\}$	•		•
SK model	as above with $\lvert I\rvert = 2$	•		•
Nk landscapes	see [188]	∘	•	∘
Graph bipartitioning	$\vartheta_{ij}([A,B]) = \begin{cases} 1 & \text{if } \{i,j\} \not\subseteq A, B \\ 0 & \text{otherwise} \end{cases} \quad i<j$	∘	•	∘
Asymmetric TSP	$\vartheta_{kl}(\tau) = \sum_i \delta_{k,\tau(i)}\delta_{l,\tau(i-1)} \qquad k \neq l$	∘	•	∘

The matrix $\mathbf{C}$ is obviously symmetric and nonnegative definite. Taking the set of all maps $\{f : X \to \mathbb{R}\}$ as base space of the probability space Ω, a basis is formed by the set of orthonormal eigenvectors $\{\psi_k\}$ of the covariance matrix $\mathbf{C}$. An expansion of the form

$$f(x) \doteq \sum_k b_k \psi_k(x) \qquad \text{a.s. } x \in X \tag{3.14}$$

is known as the *Karhunen–Loève series* or the *principal component decomposition* of f. A crucial result is the following.

THEOREM 3.6 (see [115, 122, 144]). *Let σ_k^2 denote the eigenvalue of $\mathbf{C}$ belonging to the eigenvector ψ_k. Then the coefficients of the Karhunen–Loève series* (3.14) *are uncorrelated random variables satisfying*

$$\operatorname{Cov}[b_k, b_j] = \sigma_k^2 \delta_{kj}, \qquad 1 \le k, l \le |X|. \tag{3.15}$$

Thus $\sigma_k^2 = \mathbb{V}[b_k]$. Furthermore we have $\sigma^2 = \operatorname{Tr}\mathbf{C} = \sum_k \sigma_k^2$.

Random landscapes of practical importance often exhibit strong regularities.

DEFINITION 3.7. *A random landscape $\mathcal{F}$ is* pseudoisotropic *if there are constants a_0, v, and w such that for all $x \in X$,*

(i) $\mathbb{E}[f(x)] = a_0$,
(ii) $\mathbb{V}[f(x)] = v^2$, *and*
(iii) $|X|^{-1} \sum_{y\in X} \mathbf{C}_{xy} = w$.

3.2.1. Additive Landscapes. Many important random landscapes can be written as a sum of components with random coefficients; see Table 3.2. More precisely, let M be a finite index set, let c_j, $j \in M$, be independent, real-valued random variables over appropriate probability spaces $\Omega_j = (\mathbb{R}, \mathcal{A}_j, \mu_j)$, and let $\Theta = \{\vartheta_j : X \to \mathbb{R} \mid j \in M\}$ be a family of real-valued functions on X. An *additive random landscape* (arl) is the probability space $(\mathbf{\Omega}_X, \otimes_j \mathcal{A}_j, \otimes_j, \mu_j)$ with

$$\mathbf{\Omega}_X = \left\{ f : X \longrightarrow \mathbb{R} \mid f(x) = \sum_{j=1}^{M} c_j \vartheta_j(x) \right\}. \tag{3.16}$$

In other words, the random landscape is constructed as a linear combination of non-random landscapes ϑ_j with independent random coefficients c_j.

Using the Karhunen–Loève decomposition, (3.14), any random landscape can be written as a linear combination with uncorrelated random coefficients. Since uncorrelated Gaussian random variables are independent we have the following.

LEMMA 3.8 (see [188]). *Every Gaussian random landscape is additive.*

The most important additive random landscapes exhibit additional regularities.

DEFINITION 3.9. *An arl is* uniform *if and only if*

(i) *the random variables c_i, $i \in M$, are independent, identically distributed (i.i.d.) and*

(ii) *there exist constants $a, b \in \mathbb{R}$ such that $\sum_{x\in X} \vartheta_i(x) = |X|a$ and $\sum_{x\in X} \vartheta_i^2(x) = |X|b$ holds.*

A uniform random landscape is strictly uniform *if for any $x \in X$ there exist constants $d, e \in \mathbb{R}$ with $\sum_j \vartheta_i(x) = d$ and $\sum_j \vartheta_i^2(x) = e$.*

THEOREM 3.10 (see [188]). *A uniform random landscape is pseudo-isotropic if and only if* (i) *$\mathcal{F}$ is strictly uniform, or* (ii) *$a = 0$, $\mathbb{E}[c_i] = 0$, and there is a constant $e \in \mathbb{R}$ such that $\sum_i \vartheta_i^2(x) = e$ for all $x \in X$.*

3.2.2. Isotropy. Uniformity and pseudo-isotropy are still rather weak properties. In [210, 216] the notion of an *isotropic* random landscape was introduced as a "statistically symmetric model," that is, as a random landscape with a covariance matrix that shares the symmetries of the underlying configuration space.

DEFINITION 3.11. *A random landscape is* isotropic w.r.t. a partition $\mathcal{R}$ *of $X \times X$ if there are constants a_0 and s and a function $c : \mathcal{R} \to \mathbb{R}$ such that*

(i) *$\mathbb{E}[f(x)] = a_0$ and $\mathbb{V}[f(x)] = s^2$ for all $x \in X$, and*

(ii) *$\mathbf{C}_{xy} = c(\mu)$ for all $(x, y) \in \mu$, i.e., the covariance matrix $\mathbf{C}$ is constant on the classes $\mu \in \mathcal{R}$.*

The notion of isotropy for random landscapes is the analogue of *stationarity* for stochastic processes. Following the conventions of Karlin and Taylor [123] our notion of isotropy would be called "covariance isotropic," "weakly isotropic," or "wide sense isotropic." For a Gaussian random landscape the notions of (weak) isotropy and strict isotropy coincide of course.

A partition $\mathcal{R}$ of $X \times X$ is *homogeneous* if the diagonal $\{(x, x)|x \in X\}$ is a class of $\mathcal{R}$. It is *class degree regular* if for a given class $\mathcal{X} \in \mathcal{R}$ the number $|\{y \in X|(x, y) \in \mathcal{X}\}|$ is independent of $x \in X$.

THEOREM 3.12 (see [216]).

(i) *If $\mathcal{F}$ is isotropic w.r.t. a homogeneous class degree regular partition of $X \times X$, then $\mathcal{F}$ is pseudo-isotropic.*

(ii) *If $\mathcal{F}$ is isotropic w.r.t. a homogeneous class degree regular partition of $X \times X$ and $\mathbb{E}[f(x)] = a_0$ for all $x \in X$, then $\mathcal{F}$ is isotropic w.r.t. a homogeneous coherent configuration if and only if $\mathbf{C} \in \langle\langle \mathcal{R} \rangle\rangle$.*

If $\mathbf{A}$ is the adjacency matrix of an undirected graph (or, more generally, the symmetric transition matrix of a Markov process on X), then we say that a random landscape is **-isotropic w.r.t.* $\mathbf{A}$ if $\mathbb{E}[f(x)] = a_0$ and $\mathbf{C} \in \langle \mathbf{A} \rangle$, i.e., if $\mathbf{C}$ can be written as a polynomial of $\mathbf{A}$. For association schemes (such as those arising from distance regular graphs including the hypercube) isotropy and *-isotropy are equivalent.

THEOREM 3.13 (see [216]). *An arl is *-isotropic if and only the Fourier coefficients (w.r.t. an orthonormal basis of eigenvectors of $\mathbf{A}$) satisfy the following:*

(i) *$\mathbb{E}[a_k] = 0$ for $k \neq 0$,*

(ii) *$\mathrm{Cov}[a_k, a_j] = \delta_{kj}\mathbb{V}[a_k]$, and*

(iii) *$\mathbb{V}[a_k] = \mathbb{V}[a_j]$ if φ_j and φ_k belong to the same eigenspace.*

These conditions mean that the Fourier coefficients are uncorrelated and that they have the same mean and variance whenever they belong to the same mode (eigenspace of $\mathbf{A}$). One might also say that Fourier and Karhunen–Loève series coincide for *-isotropic landscapes.

The class of *-isotropic models (on their natural configuration spaces) includes among others Derrida's p-spin Hamiltonians, the graph-bipartitioning problem, and the TSP. On the other hand, most variants of Kauffman's Nk-model, the XY-Hamiltonians, short-range Ising models, or the graph-matching problem are not isotropic. This has important implications for the structure of these landscapes. In the case of the Nk-models, for instance, it means that different positions along the sequence have a different average influence on the fitness, i.e., the expected effect of mutations may depend explicitly on the sequence position.

3.2.3. Entropy. For a random landscape with measure μ we define the *entropy*

$$S = -\int \ln \mu(f) d\mu(f). \tag{3.17}$$

It is well known that the Gaussian distributions maximize entropy. The proof for the one-dimensional case can be found, e.g., in [105, Prop. 1.15]. A proof of the general case is provided in [214] together with a more detailed discussion of the relationships between isotropy and entropy. The main result follows.

THEOREM 3.14. *Let $\mathcal{F}$ be a random landscape with positive definite covariance matrix $\mathbf{C}$ with eigenvalues $\sigma_k^2 > 0$. Then the entropy satisfies*

$$S \leq S_{\mathbf{C}} = \frac{1}{2}|X| \ln \frac{2\pi e}{|X|} + \frac{1}{2}\sum_k \ln \frac{\Lambda_k |X|}{\sigma^2}. \tag{3.18}$$

It is easy to verify that $S_{\mathbf{C}}$ is the entropy of a Gaussian distribution with covariance matrix $\mathbf{C}$.

The two terms in (3.18) allow for a direct interpretation. The Gaussian entropy $S_{\mathbf{C}}$ attains its maximum subject to a given variance σ^2 if and only if $\Lambda_k = \sigma^2/|X|$, in which case the second term vanishes. We therefore split the entropy of a random landscape into three contributions,

$$S = S_{\sigma^2} + \Delta S_{\mathbf{C}} + \Delta S_{ng}, \tag{3.19}$$

where $\Delta S_{ng} = S - S_{\mathbf{C}}$ is the entropy loss due to deviations from a Gaussian distribution, S_{σ^2} is the maximal entropy with given variance σ^2, and $\Delta S_{\mathbf{C}}$, the second term in (3.18), measures the entropy loss due to variations in the spectrum of $\mathbf{C}$. In particular, whenever there are correlations between different vertices, then $\mathbf{C}$ is nondiagonal and hence $\Delta S_{\mathbf{C}} < 0$. More precisely, $\Delta S_{\mathbf{C}} = 0$ if and only if the corresponding Gaussian random landscape is i.i.d. A Gaussian random landscape is *-isotropic if and only if ΔS_C is maximal subject to given values of $\mathbb{V}[a_k]$ in Theorem 3.13. In this sense we can regard *-isotropy as a "maximum entropy" condition.

3.3. Amplitude Spectra. Equation (3.4) decomposes nonelementary landscapes in a natural way into a superposition of elementary ones. A natural way of quantifying this decomposition is to consider the projection f_Λ of the landscape f onto the eigenspace of $-\boldsymbol{\Delta}$ with eigenvalue Λ. The relative importance of the Λ eigenspace is then quantified by the ratio of the landscape variances of f_Λ and f,

$$B(\Lambda) = \frac{\langle \tilde{f}_\Lambda, \tilde{f}_\Lambda \rangle}{\langle \tilde{f}, \tilde{f} \rangle}, \tag{3.20}$$

where, as usual, $\tilde{f}(x) = f(x) - \overline{f}$. We call $B(\Lambda)$ the *amplitude* of (the eigenspace associated with) Λ. In terms of the Fourier decomposition we obtain immediately

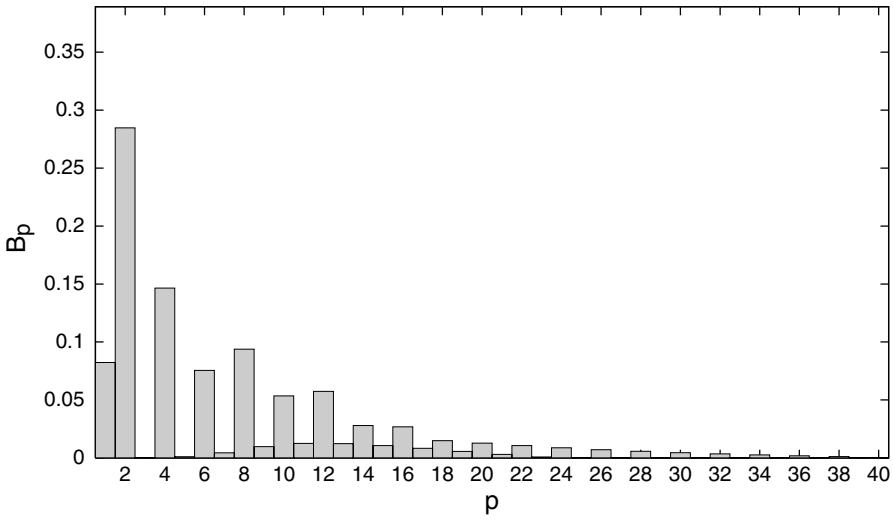


Fig. 3.1 *The estimated amplitude spectrum for a GC landscape with $n = 100$ under mutation* [114]. *The configuration space is the Hamming graph $\mathcal{Q}_2^{100}$ of sequences taken from the* 2*-letter alphabet $\{G,C\}$. The most striking feature of the amplitude spectrum of RNA landscapes is a strong difference between even and odd modes. This can easily be explained in terms of the physics underlying RNA folding: The major contribution of the folding energy comes from stacking of base pairs. Hence the major changes in free energy caused by a point mutation will arise from these contributions. Stacking energies are influenced by an even number of nucleotides depending on the location of the affected base pair within a stack. A recent comparison of amplitude spectra for different landscapes based on folding short RNA chains indicates that the amplitude spectra of the free energy landscapes are typical* [212].

(see [211])

$$B(\Lambda) = \sum_{k|-\boldsymbol{\Delta}\varphi_k=\Lambda\varphi_k} |a_k|^2 \Big/ \sum_{k|-\boldsymbol{\Delta}\varphi_k\neq 0} |a_k|^2. \tag{3.21}$$

For convenience of notation we set $B(0) = 0$. Thus $B(\Lambda) \geq 0$ and $\sum_\Lambda B(\Lambda) = 1$.

In many cases, in particular for landscapes on Hamming graphs, it is more convenient to refer to an amplitude by the interaction order (number of the eigenvalue when eigenvalues are arranged in ascending order *without* counting multiplicities). Hence one typically finds B_p instead of $B(p\alpha)$ for the pth eigenspace of an α-letter Hamming graph. Obviously, the Laplacian $-\boldsymbol{\Delta}$ can be replaced, for instance, by a transition operator if desired. In the case of random landscapes one naturally considers the expectation values $\mathbb{E}[B(\Lambda)]$.

Amplitude spectra are a very useful way of classifying nonelementary landscapes. We mention just a few examples here.

- Asymmetric TSPs are superpositions of symmetric and antisymmetric TSPs with equal weights [211]. Together with Table 3.1 this observation explains the behavior of the TSP mentioned in the introduction: with transpositions the symmetric and the antisymmetric versions of the TSP have very similar Laplacian eigenvalues and hence essentially the same correlation structure. In the case of reversals, however, we have a smooth landscape belonging to the third eigenvalue for symmetric C, and an essentially uncorrelated landscape corresponding to a highly "excited state" for the antisymmetric component.

- Assignment problems also have permutations as their underlying configurations [27, 151]. The quadratic assignment problem consists in general of three contributions corresponding to the three smallest nonzero eigenvalues of the Laplacian of the Cayley graph $\Gamma(\mathsf{S}_n, \mathcal{T})$ [192].
- The landscape of the *low autocorrelated binary string problem* consists of a dominating 4-spin contribution and an asymptotically vanishing 2nd-order component [211].
- The free energy landscapes of 2-letter RNA sequences consist predominantly of the small even modes $p = 2, 4, 6, 8$ [114], while the biophysically more relevant 4-letter alphabet AUGC has an additional large $p = 1$ component; see Figure 3.1. See section 5.1.1 for an explanation of the RNA model.
- Widely used methods for reconstructing phylogenies from molecular data, such as maximum parsimony and maximum likelihood approaches [57, 58], require the maximization of a score function on the set of all leaf-labeled trees; see Figure 2.1. Amplitude spectra show that these landscapes are highly correlated [17].

Amplitude spectra of landscapes arising from more complicated models, such as evolving cellular automata [40], RNA folding, or electronic circuit design have been computed as well [114, 192, 232].

4. Ruggedness and Neutrality. Ruggedness intuitively is just the opposite of smoothness. Neutrality, i.e., the existence of neighboring configurations with the same fitness, appears to be just a way of achieving a "smooth" landscape. Therefore it comes as a surprise that ruggedness and neutrality turn out to be independently tunable parameters.

4.1. Ruggedness and Autocorrelation Functions. The ruggedness of a landscape is most easily quantified by measuring the correlation of fitness values in "neighboring" positions. Weinberger [242, 243] suggested the following procedure. Given a Markov process on X, we sample the fitness values $f(x^{(t)})$, interpret them as a time series, and compute the autocorrelation function of this time series. Let $\mathbf{T}$ be the transition matrix of such a reversible Markov process with stationary distribution φ_0. The (expected) autocorrelation function along a $\mathbf{T}$-random walk on X is then

$$r(t) = \left(\sum_{x \in X} |\tilde{f}^2(x)| \varphi_0(x) \right)^{-1} \sum_{y \in X} \tilde{f}(x) (\mathbf{T}^t)_{xy} \tilde{f}^*(y) \varphi_0(y) = \frac{\langle \tilde{f}, \mathbf{T}^t \tilde{f} \rangle_{\varphi_0}}{\langle \tilde{f}, \tilde{f} \rangle_{\varphi_0}}. \tag{4.1}$$

By expanding f w.r.t. eigenvectors of $\mathbf{T}$ it can be shown [211] that

$$r(t) = \sum_{\lambda \neq 1} B_{\mathbf{T}}(\lambda) \lambda^t, \tag{4.2}$$

where $B_{\mathbf{T}}(\lambda)$ are the amplitudes of f w.r.t. the eigenspaces of $\mathbf{T}$. Thus a landscape f is elementary w.r.t. a transition operator $\mathbf{T}$ if and only if the "random walk" autocorrelation function is exponential, $r(t) = \lambda_p^t$.

For regular graphs $\mathbf{T}$, $\mathbf{A}$ and $-\mathbf{\Delta}$ have the same eigenspaces, and the eigenvalues of the transition matrix can be expressed in terms of the Laplacian eigenvalues as $\lambda = 1 - \Lambda/D$, where D is the vertex degree of the graph. Thus (4.2) becomes

$$r(t) = \sum_{\Lambda \neq 0} B(\Lambda) \left(1 - \Lambda/D\right)^t. \tag{4.3}$$

The information contained in $r(s)$ is therefore equivalent to the amplitude spectrum. A landscape is highly correlated if $r(s)$ decays slowly, i.e., if $B(\Lambda)$ is large for small eigenvalues Λ. The *correlation length*

$$\ell = \sum_{t=0}^{\infty} r(t) = D \sum_{\Lambda \neq 0} \frac{B(\Lambda)}{\Lambda} \tag{4.4}$$

may be used to condense the correlation information into a single measure of ruggedness.

Most of the early work on RNA landscapes, e.g., [73, 225], used a different type of correlation measure based on the Hamming distance. In [211, 216] this approach was generalized to relations on $\mathcal{R}$ on $X \times X$.

DEFINITION 4.1. *Given a relation $\mathcal{R}$ on $X \times X$, the autocorrelation of f w.r.t. $\mathcal{R}$ is*

$$\varrho(\mathcal{R}) = \frac{|X|^2}{|\mathcal{R}|} \frac{\sum_{(x,y)\in\mathcal{R}} (f(x) - \overline{f})(f(y) - \overline{f})}{\sum_{x,y\in X} (f(x) - \overline{f})(f(y) - \overline{f})}. \tag{4.5}$$

On Hamming graphs, for instance, it is natural to consider the distance classes, i.e., $(x, y) \in \mathcal{R}_d$ if and only if $d_H(x, y) = d$. Such distance-dependent correlation functions have been considered also for some combinatorial optimization problems [4, 5, 204, 215]. Given a partition of $X \times X$, we may of course regard ϱ as a function of the classes of this partition. Furthermore, if this partition is sufficiently "nice," then the correlation function ϱ itself also has useful algebraic properties. An example is the following.

THEOREM 4.2 (see [209]). *Let f be a landscape on a graph Γ that has a homogeneous coherent algebra $\mathfrak{W}[\Gamma]$. Then $r(s)$ is exponential if and only if ϱ is a left eigenvector of the collapsed adjacency matrix $\hat{\mathbf{A}}$.*

In practice $\varrho(d)$ and $r(t)$ convey the same information. Explicit expressions relating these functions with each other can be found in [211]. In particular, we have $\varrho(1) = r(1)$ for the nearest neighbor correlation, which is the most commonly used parameter of ruggedness.

4.2. Ruggedness and Local Optima.

4.2.1. The Number of Local Optima. Local optima play an important role since they might be obstacles on the way to the optimal solution. In the theory of disordered systems, local minima of the energy function are usually called metastable states. For the sake of definiteness we shall consider *local minima*, i.e., configurations $\hat{x} \in X$ satisfying

$$f(x) \leq f(y) \quad \forall\, y \in N(x)\,. \tag{4.6}$$

Analogous expressions for local maxima can be obtained by replacing f with $-f$. The number and distribution of local minima provide an alternative approach to landscape ruggedness.

In [173] Palmer proposed calling a landscape f *rugged* if the number M_f of local optima scales exponentially with some measure of "system size" such as the number of cities in a TSP or the number of spins in spin glasses. Unfortunately, there is in general no simple way of computing M_f without exhaustively generating the landscape. Alternatively, one can of course estimate M_f by checking whether a randomly

generated $x \in X$ is a local minimum. Numerical data of this kind are reported, e.g., in [81, 218, 217]. Methods from statistical mechanics can also be used to obtain the scaling of the expected value $\mathbb{E}[\mathsf{M}]$ with the system size for a variety of disordered systems; see, e.g., [24, 43, 45, 59, 93, 190, 226, 228].

A nonrigorous result is of particular interest in this context. The *correlation length conjecture* [218] suggests that the number of local optima of a "typical" landscape can be estimated from its correlation length ℓ, (4.4). More precisely, one expects on the order of one local optimum on a mountain with a radius that is determined by the correlation length ℓ. Numerical surveys provided good evidence that the correlation length conjecture yields a fairly accurate prediction of the number of local optima (metastable states) of isotropic elementary random landscapes; see [81] for a summary of the numerical data.

4.2.2. Basins. To each local minimum $\hat{x}$ there is an associated *basin* $\mathcal{B}(\hat{x})$ defined by means of the steepest descent algorithm: Starting with $z_0 = y$ we choose at each step the neighbor $z_{k+1} \in N(z_k)$, $f(z_{k+1}) < f(z_k)$ with the smallest fitness value and repeat the procedure until it terminates when $z_{k+1} = \hat{x}$ is a local minimum. The notion of a basin hence may become ambiguous when there is "local neutrality" in $N(x)$, i.e., if there are $x \in X$ and $y, y' \in N(x)$ with $f(y) = f(y')$. It is an open question how the basin should be defined in full generality. It is not surprising that the distribution of basin sizes is crucial for the performance of simple optimization heuristics.

Probably the simplest approach is to generate an initial configuration at random and then to use steepest descent to reach the minimum of the basin. The question then becomes how likely it is to hit the basin of the global optimum by chance. Let α_j, $j = 1, \ldots, \mathsf{M}_f$, denote the sequence of relative basin sizes $|\mathcal{B}(\hat{x})|/|X|$.

THEOREM 4.3 (see [83]). *The probability $p(m)$ that in a sample of m randomly chosen configurations we find at least one configuration in each basin is*

$$p(m) = \sum_{k=0}^{\mathsf{M}_f} (-1)^{\mathsf{M}_f - k} \sum_{1 \le j_1 \le \cdots \le j_k \le \mathsf{M}_f} (\alpha_{j_1} + \cdots + \alpha_{j_k})^m. \tag{4.7}$$

From this rather complicated expression one can deduce, for instance, the following corollary.

COROLLARY 4.4 (see [83]). *Suppose $\mathsf{M}_f \gg 1$, $m = a^2 \mathsf{M}_f$ for some $a > 0$, and the relative basin sizes are uniformly distributed. Then $p(m) = \exp(-1/a)$.*

In other words, sampling $\mathcal{O}(\mathsf{M}_f^2)$ points at random provides a finite chance to find the basin of the global optimum.

However, so far there does not appear to be a good method for estimating basin sizes beyond exhaustive enumeration or random sampling. An important aspect is the correlation between basin size and fitness of the minimum: In highly correlated landscapes, i.e., landscapes in which the amplitude spectrum shows large values of $B(\Lambda)$ for small values of Λ only, it appears that deeper minima have larger basins [59, 213].

4.2.3. Barriers and Depth. The basins of local minima are separated by saddle points and fitness barriers. Let $\hat{x}$ and $\hat{y}$ be two local minima and let $\mathbf{p}$ be a path in X from $\hat{x}$ to $\hat{y}$. Then the fitness barrier separating $\hat{x}$ from $\hat{y}$ is

$$f[\hat{x}, \hat{y}] = \min \left\{ \max \left[f(z) \middle| z \in \mathbf{p} \right] \;\middle|\; \mathbf{p} : \text{path from } \hat{x} \text{ to } \hat{y} \right\}. \tag{4.8}$$

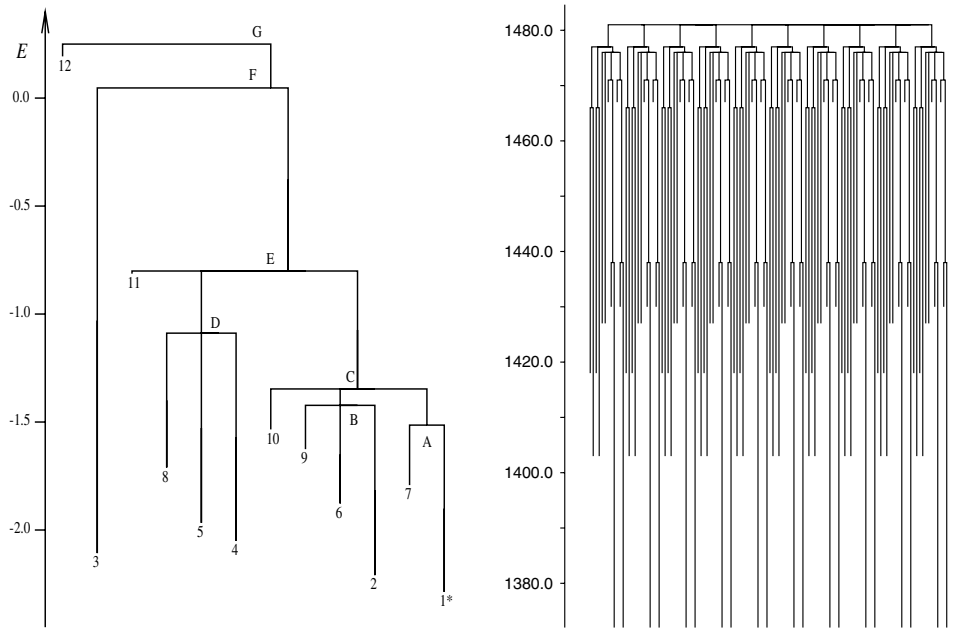


Fig. 4.1 *Barrier trees. A guide to reading the trees is given on the left using a simple example (Gaussian random numbers on a $\mathcal{Q}_2^7$): The leaves* 1–12 *denote the local minima. The global minimum* 1 *is marked with an asterisk. Saddle points are labeled with capital letters from A to G. The saddle points* B, C, D, E *are "degenerate," indicating that the lowest energy paths leaving, e.g.,* 4,5,8 *run through a common exit point. Note that all* $2^7 = 128$ *configurations have pairwise distinct energies, hence there are no two distinct saddle points with the same energy, which may exist in general. The barrier of* 5 *is* $B(5) = E(\mathsf{D}) - E(5)$ *along the lowest path from* 5 *to* 4*, while* $B(4) = E(\mathsf{E}) - E(4)$ *along the lowest path from* 4 *to* 1^**. The tree on the right shows all local minima and their barriers for the "New Mexico weekend trip TSP" (Figure* 1.3*) with transpositions as move set. The high symmetry reflects the fact that the length of symmetric TSP tours does not change when the starting point is shifted along the tour and when the tour is traveled in reverse order.*

A point $\hat{z} \in X$ satisfying the minimax condition in (4.8) is a *saddle point* of the landscape. The saddle-point energies $f[\hat{x}, \hat{y}]$ form an ultrametric distance measure on the set of local minima; see, e.g., [160, 177, 233]. This hierarchical structure can be represented by the *barrier tree* of the landscape [18, 64, 67, 84, 163, 239]; see Figure 4.1. Its leaves are the local minima and its internal nodes correspond to saddle points.

The *barrier* enclosing a local minimum is the height of the lowest saddle point that gives access to a more favorable minimum. In symbols,

$$B(\hat{x}) = \min\left\{ f[\hat{x}, \hat{y}] - f(\hat{x}) \middle| \hat{y} : f(\hat{y}) < f(\hat{x}) \right\}. \tag{4.9}$$

If $B(\hat{x}) = 0$, then the local minimum $\hat{x}$ is degenerate. It is easy to check that (4.9) is equivalent to the definition of the depth of a local minimum in [128]. For metastable states it agrees with the more general definition of the depth of a "cycle" in the literature on inhomogeneous Markov chains [8, 30, 31]. The information contained in the energy barriers is conveniently summarized by two global parameters. Let Ω_f be the set of all global minima of f:

$$\mathsf{D} = \max\left\{ B(s) \middle| s \notin \Omega_f \right\}, \tag{4.10}$$

$$\psi = \max\left\{ \frac{B(s)}{f(s) - f(\min)} \middle| s \notin \Omega_f \right\}. \tag{4.11}$$

Both parameters are easily obtained from the barrier tree. The *depth* D and *difficulty* ψ [30, 31, 98, 128, 195] play a crucial role in the theory of simulated annealing; see section 6.1.

4.3. Neutrality. In this section we consider exclusively landscapes on finite simple undirected graphs with vertex set X and edge-set E. The number of neutral neighbors of $x \in X$ is

$$\nu(x) = \sum_{y \in N(x)} \delta\big(f(x), f(y)\big), \tag{4.12}$$

which can again be regarded as a landscape on Γ. Empirically, it turns out that the important classes of landscapes arising in evolutionary biology and the theory of computer simulations exhibit a large degree of neutrality. We postpone a discussion of these cases until section 5. Below we briefly discuss two mechanisms for generating neutrality in simple models.

4.3.1. Neutrality from Symmetry. A rather trivial cause of neutrality are symmetries in the fitness function. In some cases these symmetries arise from embedding a combinatorial optimization problem in a state space that is too large. We briefly outline one example.

Given a weighted graph with an even number n of vertices and a symmetric weight matrix $\mathbf{W}$, the task of the *graph-matching problem* (GMP) is to determine a set ξ of $n/2$ edges (i_k, j_k) with distinct incident vertices (a matching) such that

$$f(\xi) = \sum_{k=1}^{n/2} \mathbf{W}_{i_k, j_k} \tag{4.13}$$

is maximized. A matching ξ is conveniently encoded as a permutation π such that $\{\pi(2k-1), \pi(2k)\}$ is an edge of the matching. The resulting landscape is elementary on the Cayley graph $\Gamma(\mathsf{S}_n, \mathcal{T})$ with the transpositions as generators. Obviously, canonical transpositions of the form $\tau_{2k-1,2k}$ leave $f(\xi)$ unchanged since they merely exchange the endpoints of the same edge and hence lead to the same matching.

4.3.2. Neutrality in Additive Random Landscapes. At present the most interesting approaches to neutrality focus on random landscapes. Hence we shall consider the random variable version of (4.12), as follows.

DEFINITION 4.5. *Let $\mathcal{F}$ be a random landscape. The random variable*

$$\nu_x : \mathcal{F} \longrightarrow \mathbb{Z}, \quad \nu_x(f) = \sum_{x' \in N(x)} \delta\big(f(x), f(x')\big) \tag{4.14}$$

is called the neutrality of $\mathcal{F}$ in x.

It is convenient to define the following parameters for all $y, y', y'' \in N(x)$:

$$c_x(y) = \big|\{j \in M \mid \vartheta_j(x) \neq \vartheta_j(y)\}\big|, \tag{4.15}$$

$$w_x(y', y'') = \big|\{j \in M \mid \vartheta_j(x) \neq \vartheta_j(y') \ \wedge \vartheta_j(x) \neq \vartheta_j(y'')\}\big|, \tag{4.16}$$

$$\Xi = \mathbb{E}\left[\frac{1}{|X|} \sum_x \left(\nu_x - \frac{1}{|X|} \sum_{x'} \nu_{x'} \right)^2 \right], \tag{4.17}$$

where $x \in X$ is an arbitrary vertex. The quantity Ξ is the expected variance of the family ν_x across a given landscape.

Theorem 4.6 below describes quite completely how additive random landscapes behave when their coefficients c_j vanish with a nonzero probability. This class of random landscapes is, so far, the only one for which a detailed analysis of neutrality is available. Newman and Engelhardt [165] and [12] considered variations of Kauffman's Nk landscape with integer coefficients, which also leads to nonvanishing neutrality in general.

THEOREM 4.6 (see [188]). *Let $\mathcal{F}$ be an arl with coefficients c_i having the property*

$$\mu = \mu_0 \cdot \epsilon_0 + (1-\mu_0) \cdot \nu_{\alpha,\sigma^2}, \tag{4.18}$$

where ϵ_0 is the measure concentrated in 0, ν_{α,σ^2} is the Gaussian measure with mean α and variance σ^2, and μ_0 is some real number with $\mu_0 \in [0,1]$. Then we have for finite sets $N(x)$

$$\mathbb{E}[\nu_x] = \sum_{y \in N(x)} \mu_0^{c_x(y)}, \tag{4.19}$$

$$\mathbb{V}[\nu_x] = \sum_{y',y''} \mu_0^{c_x(y')+c_x(y'')} \left[\mu_0^{-w_x(y',y'')} - 1\right], \tag{4.20}$$

$$\Xi = \frac{1}{|X|}\left[\sum_y \mathbb{V}(\nu_y) - \frac{1}{|X|}\sum_{y,y'} \mathrm{Cov}(\nu_y,\nu_{y'})\right] + \frac{1}{|X|}\sum_y \mathbb{E}[\nu_y]^2 - \left(\frac{1}{|X|}\sum_y \mathbb{E}[\nu_y]\right)^2, \tag{4.21}$$

where $\frac{1}{|X|}\sum_{y,y'} \mathrm{Cov}(\nu_y,\nu_{y'}) \geq 0$.

The reason for considering measures μ with the property (4.18) is the following: let M be a finite index set and η_j, $j \in M$, be some set of nonzero real constants. Then we have for independent random variables c_j, $j \in M$,

$$\mu \text{ has property (4.18)} \quad \Longrightarrow \quad \mu\left\{\sum_{j\in M} c_j\eta_j = 0\right\} = \mu_0^{|M|}.$$

In other words, in order to have $f(x) = f(y)$ with $y \in N(x)$ exactly $c_x(y)$ (see (4.15)) coefficients c_j have to be zero. In the case of measures of the form $\mu_0 \cdot \epsilon_\xi + (1-\mu_0)\cdot\nu_{\alpha,\sigma^2}$ for $\xi > 0$ it is possible to construct and analyze neutral landscapes in a somewhat analogous way.

In [188] a number of applications of Theorem 4.6 are discussed. Here we restrict ourselves to the simplest one, as follows.

COROLLARY 4.7. *For a p-spin model with coefficients c_i satisfying* (4.18) *we have*

$$\mathbb{E}[\nu] = n\,\mu_0^{\binom{n-1}{p-1}}, \tag{4.22}$$

$$\mathbb{V}[\nu] = n(n-1)\mu_0^{2\binom{n-1}{p-1}}\left[\mu_0^{-\binom{n-2}{p-2}} - 1\right] + n\,\mu_0^{\binom{n-1}{p-1}}\left[1 - \mu_0^{\binom{n-1}{p-1}}\right], \tag{4.23}$$

$$\Xi = 0. \tag{4.24}$$

Consider a spin-glass model where the spins are arranged on a finite-dimensional lattice. That is, independent of the size of the system, there is only a finite number of lattice neighbors for each spin. In *short-range spin glasses*, the only nonzero interaction coefficients link lattice neighbors, i.e., all but $\mathcal{O}(n)$ coefficients vanish. A short-range spin glass is therefore characterized by $\mu_0 = 1 - \frac{z}{n^{p-1}}$, where $z > 0$ is a parameter determined by the connectivity of the lattice. As a consequence we have for every short-range spin glass

$$\lim_{n\to\infty} \mathbb{E}[\nu/n] = \mathrm{e}^{-z} \quad \text{and} \quad \lim_{n\to\infty} \mathbb{V}[\nu/n] = 0. \tag{4.25}$$

The p-spin models are elementary w.r.t. spin-flip moves; see Table 3.1. On the other hand, we may use μ_0 to tune the degree of neutrality to any desired value. Conversely, given a value of $\mathbb{E}[\nu]$, we may choose p arbitrarily, thereby prescribing any desired degree of ruggedness. Thus we have established that *ruggedness and neutrality are independent features of (random) landscapes.*

5. Landscapes and Genotype-Phenotype Maps.

5.1. General Considerations. In the context of RNA sequences, fitness often does not depend on the particular sequence of nucleotides but on its actual (spatial) structure. That is, there exists a generic partition on the configuration space, the elements of which are called *phenotypes*, representing classes of genotypes. Note, however, that two phenotypes do not necessarily have different fitness values. Accordingly, we can decompose the landscape as follows:

$$\textbf{Genotypes} \longrightarrow \textbf{Phenotypes} \longrightarrow \textbf{Fitness}. \tag{5.1}$$

Obviously, many properties of the fitness landscape are closely related to properties of the genotype-phenotype mapping. We will study these connections in the following using ribonucleic acids (RNA) and sequential dynamical systems (SDSs) as paradigms. We will call the preimage of a fixed phenotype its *neutral network*[1] and we will discuss the properties and role of neutral networks. In the following we will provide some background on RNA and SDS.

5.1.1. RNA. RNA acts in viruses and cells as a messenger (mRNA), carrying the genetic information from the DNA to the translation apparatus, and as transfer RNA, or tRNA for short. It plays the role of an adapter for the synthesis of proteins and finally as ribosomal RNA (rRNA) it is an integral part of the ribosome and exhibits catalytic activities in natural polypeptide synthesis [33, 34, 246].

RNA thus serves two purposes: (i) storage of genetic information based on a one-dimensional template that can be read and copied on request, and (ii) catalytic properties as ribozymes which require three-dimensional structures in order to gain efficiency and specificity in processing specific substrates. As demonstrated by Spiegelman, *in vitro* evolution experiments can be performed to select RNA molecules that are capable of fast replication [154]. Indeed, replication rates are optimized in serial transfer experiments [54, 121, 196]. In case one wants to optimize properties other than replication, intervention is required making use of special techniques, which interfere with *natural selection*. A well-known example is represented by the SELEX method—an acronym for *systematic evolution of ligands by exponential enrichment*—which allows

[1] At this point it seems more justified to call the preimage a neutral set since it is not clear that there are any edges among the vertices. We will justify the notion *neutral net* later in the text, as it turns out that there are many edges and various connectivity properties in the preimages

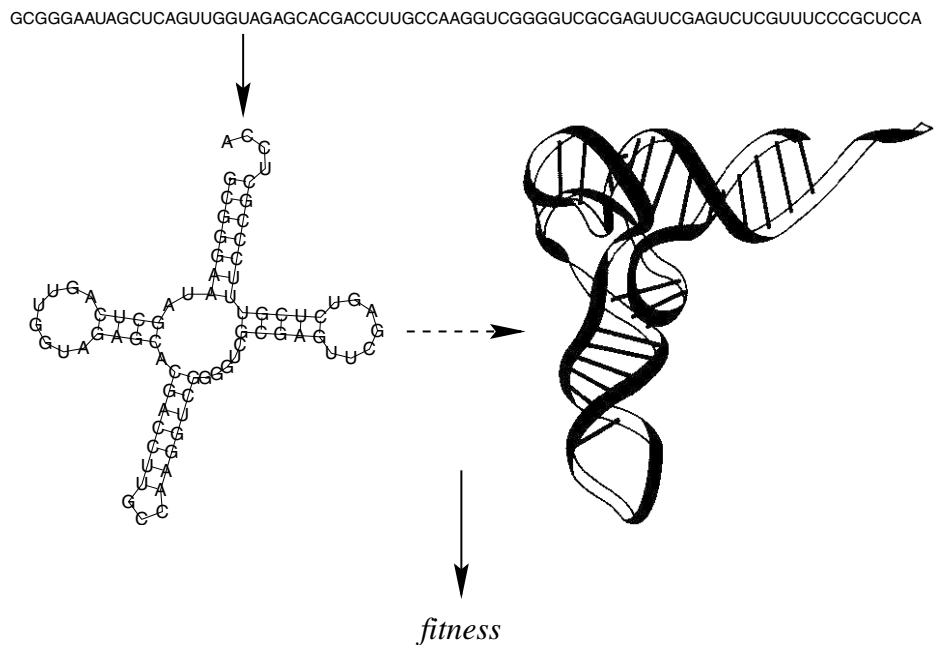


Fig. 5.1 *RNA folding landscapes. For any RNA sequence, which can be viewed as a string from the alphabet* {A, C, G, U}*, a secondary structure which minimizes energy is computed using a dynamic programming algorithm. The resulting outerplanar graph (left) is a useful approximation to the three-dimensional shape of the molecule. In computational studies the fitness of an RNA sequence is evaluated as a function of its secondary structure.*

one, for example, to create molecules with optimal binding constants [229]. The SELEX procedure is a protocol which isolates high-affinity nucleic acid ligands for a target, for example, a protein, from a pool of variant sequences. Multiple rounds of replication and selection exponentially enrich the population of species which exhibit the highest affinity, i.e., which fulfill the required task. This procedure thus allows for simultaneous screening of highly diverse pools of nucleic acid molecules for different functionalities (for a review see, e.g., [53, 133]). Results from those experiments clearly demonstrate the essential property of RNA molecules: that genotype, i.e., the RNA sequence, and phenotype, associated with the structure, are combined in one molecule.

Computer models of an RNA toy world based on the explicit computation of secondary structures (see Figure 5.1) were pioneered by the Vienna group [72, 118, 189, 199, 200, 201] and led to the first realistic models of biological landscapes; see section 5.2. For a recent review of various aspects of the RNA world *in silico*, see [66].

5.1.2. Sequential Dynamical Systems. SDSs are discrete dynamical systems that were introduced to capture basic features of computer simulations [13, 14, 15, 16, 161, 185]. An SDS consists of (a) an undirected graph Y (with vertex set $\{1, \ldots, n\}$), (b) a collection of Boolean functions (F_i) that update the state of each vertex i as a function of its neighbors while leaving all other vertex states unchanged, and (c) an *update schedule* π, defining the order in which the vertices are updated. The composition of the maps F_i in the order prescribed by the update schedule π yields the SDS $[\mathfrak{F}, \pi] = \prod_{i=1}^{n} F_{\pi(i)} : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n$. Here $\mathbb{F}_2$ denotes the finite field with 2 elements.

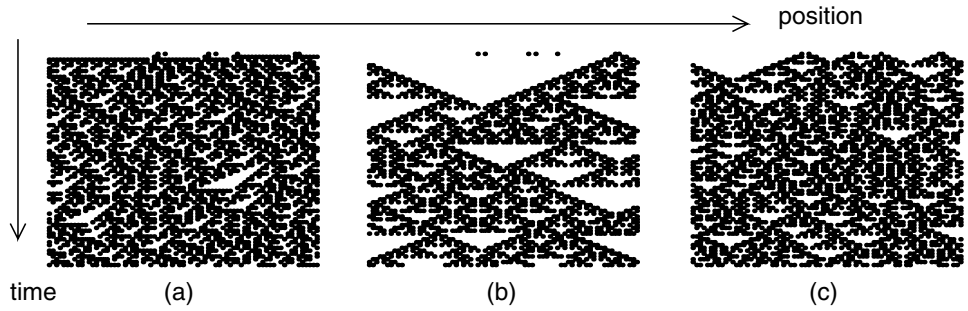
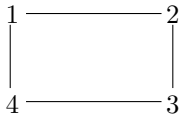


Fig. 5.2 *A particular class of SDSs are sequential cellular automata (sCA). Here we display the space time evolution of four different sCA over* Circ_{75} *with underlying rule* 90 *(applied for each vertex) and fixed initial condition. The sequential updates monitored are* (a) $(1, 2, \ldots, n)$ *(natural order),* (b) $(1, 3, \ldots, 75, 2, 4, \ldots, 74)$ *(odd-even),* (c) $(2, 4, \ldots, 74, 1, 3, \ldots, 75)$ *(even-odd).*

Example. Let $Y = \mathrm{Circ}_4$ be the circle graph on 4 vertices.



Suppose we have the parity function $p_3 : \mathbb{F}_2^3 \to \mathbb{F}_2$, $p_3(x_1, x_2, x_3) = \sum_i x_i \bmod 2$ for each vertex. Then we obtain for the update schedule $(1, 2, 3, 4)$ with initial state $(1, 1, 0, 0)$

$$\begin{aligned} F_1(1,1,0,0) &= (0,1,0,0),\\ F_2 \circ F_1(1,1,0,0) &= (0,1,0,0),\\ F_3 \circ F_2 \circ F_1(1,1,0,0) &= (0,1,1,0),\\ F_4 \circ F_3 \circ F_2 \circ F_1(1,1,0,0) &= (0,1,1,1), \end{aligned}$$

and consequently we have $[F_{\mathrm{Circ}_4}, (1, 2, 3, 4)](1, 1, 0, 0) = (0, 1, 1, 1)$.

One important problem is to analyze the set of all schedules π' that lead to the same dynamical system. In Figure 5.2 we present examples on how the update schedule affects the dynamics of an SDS.

5.2. RNA Secondary Structure Folding.

5.2.1. Genotypes and Phenotypes for RNA. In the following we will consider RNA sequences of constant length, represented by n-tuples, $(x_1, \ldots, x_n)$, with $x_i \in \mathcal{A}$, $\mathcal{A}$ being a finite alphabet formed by the nucleotides. The basic mutational mechanism consists of random point mutations that occur with independent probability. This motivates us to call two sequences adjacent if they differ in exactly one nucleotide. Sequence space with this adjacency relation is referred to as $\mathcal{Q}_\alpha^n$ (the generalized n-cube), where $\alpha = |\mathcal{A}|$. In $\mathcal{Q}_\alpha^n$ each sequence has $(\alpha - 1)n$ neighbors and the maximal distance between two sequences is n.

RNA phenotypes are in general molecular structures of various resolutions. Here we will consider the following class of RNA secondary structures [241].

DEFINITION 5.1. *A* secondary structure *over* n *vertices* $\{1, \ldots, n\}$, s_n, *is a vertex-labeled graph with an adjacency matrix* $A(s_n) = (a_{i,k})_{1 \le i,k \le n}$ *such that*

- $a_{i,i+1} = 1$ *for* $1 \leq i \leq n-1$,
- *for each* i *there is at most a single* $k \neq i-1, i+1$ *such that* $a_{i,k} = 1$,
- *if* $a_{i,j} = a_{k,l} = 1$ *and* $i < k < j$, *then* $i < l < j$.

We call an edge $\{i,k\}$, $|i-k| \neq 1$, *a base pair. A vertex* i *connected only to* $i-1$ *and* $i+1$ *is called unpaired.*

The combinatorics of secondary structures viewed as abstract graphs has been studied in detail; see, e.g., [103, 109, 116, 197]. A particular result from asymptotic combinatorics on secondary structures—with certain restrictions such as minimum helix length—is that their number asymptotically becomes $O(a^n)$ with $a < 2$ [109]. This result immediately implies that there are structures having preimages of exponential size. Moreover, the RNA model allows for several generic choices of the fitness assignment, such as, for example, using the thermodynamic stability and the degradation constant of the corresponding secondary structure.

5.2.2. The Intersection Theorem. One important question in the context of neutral networks is how close two such networks come in sequence space. In order to investigate this question we need some terminology: We call a nucleotide sequence (x_i) compatible w.r.t. a secondary structure s_n if and only if for all $a_{i,k}$ with $a_{i,k} = 1$ and $k \neq i-1, i+1$ the nucleotides x_i and x_k *could in principle* form a Watson–Crick base pair. We denote the set of compatible sequences w.r.t. some secondary structure s_n by $C(s_n)$. Note that we have

$$C(s_n) \cong \mathcal{Q}_\alpha^{n_1} \times \mathcal{Q}_\beta^{n_2},$$

where n_1, n_2 are the numbers of unpaired bases and base pairs, respectively, and β is the size of the alphabet formed by the base pairs, i.e., all pairs of nucleotides that can actually establish a chemical bond.

In terms of combinatorics, the uniqueness property of the Watson–Crick base pairs of an RNA secondary structure corresponds to an involution (viewing the base pairs as transpositions within the symmetric group, S_n [187, 189]).

THEOREM 5.2 (see [189]). *Let* s_n^1, s_n^2 *be two secondary structures with the sets of compatible sequences* $C(s_n^1)$, $C(s_n^2)$. *Then*

$$C(s_n^1) \cap C(s_n^2) \neq \varnothing. \tag{5.2}$$

Accordingly, for any two secondary structures there exists a sequence that *could* in principle realize both, from which we can conclude that the corresponding neutral networks come relatively close in sequence space. This in not true for more than two sequences. A necessary and sufficient condition for the intersection of the compatible sets of an arbitrary number of secondary structures to be nonempty can be found in [65]. An investigation of the structure of the intersections of random structures can be found in [183].

The above theorem is the basis of an experimental study [198] in which an RNA sequence is presented that simultaneously realizes two structurally dissimilar ribozymes. Of course, this sequence is taken from the intersection of the corresponding sets of compatible sequences.

5.2.3. Connectivity of Neutral Networks. In the following, we will denote a probability measure by μ_n, where n refers to some index of the corresponding probability space Ω_n (here, a random graph) over n vertices. Let P_n be some property

(event) in Ω_n. Then we write "P_n holds almost surely (a.s.)"[2] if and only if we have $\lim_{n\to\infty}\mu_n\{P_n\}=1$.

The random graph model. Let $\mathcal{Q}_\alpha^n$ be a generalized n-cube over an alphabet of length α. Let Γ_n be a subgraph of $\mathcal{Q}_\alpha^n$ and $\mu_n\{\Gamma_n\}=\lambda_n^{|\Gamma_n|}(1-\lambda_n)^{\alpha^n-|\Gamma_n|}$. Then we call $\mathcal{Q}_{\alpha,\lambda_n}^n$ the random induced subgraph model.

For RNA folding landscapes we can interpret the probability λ as the neutrality degree, i.e., the number of neutral neighbors, ν, divided by the total number of neighbors, $(\alpha-1)n$.

THEOREM 5.3 (see [179]). *In $\mathcal{Q}_{\alpha,\lambda_n}^n$, let $C_n^{(1)}$ be the largest component of a $\mathcal{Q}_\alpha^n$-subgraph Γ_n. Then there exists a constant $c>0$ such that for $\lambda_n\geq\frac{c\ln(n)}{n}$ the subgraph $C_n^{(1)}$ is a giant component, i.e.,*

$$|C_n^{(1)}|\sim|\Gamma_n| \qquad a.s.$$

holds.

It may be of interest to note that Theorem 5.3 establishes the existence of the giant component *indirectly*. The proof gives no clue on how to construct a path between two vertices and, moreover, on how long such a path might be. The explicit construction of (short) paths between vertices of neutral networks would therefore be of particular interest and leads to a deeper understanding of the likelihood that such a path would be realized in an evolutionary search. In fact the next result provides such a *constructive* proof, although we will need the higher probability $\lambda_n\geq n^{-a}$ with $0\leq a<1/2$.

Giant components of random subgraphs of $\mathcal{Q}_2^n$ obtained by independent selection of $\mathcal{Q}_2^n$-*edges* have been investigated in [1].[3] There the authors gave a sharp threshold value for the existence of a giant component at $p=1/n$. Giant components in the random graph $G_{n,p}$ exhibit the threshold value $p=1/n$ [55]. Although the actual threshold values for $G_{n,p}$ random subgraphs of $\mathcal{Q}_2^n$ and random induced subgraphs of generalized n-cubes are very similar, the corresponding methods of proof differ significantly: in the case of giant components of $G_{n,p}$ a branching process can be used whereas the above theorem and [1] are proved indirectly by showing the impossibility of keeping two "big" components separate in *two* separate randomizations.

THEOREM 5.4 (see [182]). *Let $0\leq a<1/2$ and let $k\in\mathbb{N}$ with $k>\frac{1+3a}{1-2a}$. In $\mathcal{Q}_{\alpha,\lambda_n}^n$ let λ_n be such that $\exists n_0\in\mathbb{N}$; $\forall n\geq n_0$, $\lambda_n\geq n^{-a}$ holds. Finally, let $d_{\mathcal{Q}_\alpha^n}$ and d_{Γ_n} denote the distances in the graphs $\mathcal{Q}_\alpha^n$ and Γ_n, respectively. Then*

$$\forall\, P,Q\in\mathcal{Q}_\alpha^n,\quad \lim_{n\to\infty}\mu_n\{d_{\Gamma_n}(P,Q)\leq[2k+3]\,d_{\mathcal{Q}_\alpha^n}(P,Q)\}=\lambda_n^2$$

holds. In particular for constant $\lambda_n=\lambda$ we have

$$\forall\, P,Q\in\mathcal{Q}_\alpha^n,\quad \lim_{n\to\infty}\mu_n\{d_{\Gamma_n}(P,Q)\leq 7d_{\mathcal{Q}_\alpha^n}(P,Q)\}=\lambda^2.$$

Let us finally come to the last result on generalized n-cubes. From now on we will assume a constant probability $\lambda>0$.

[2]Strictly speaking the notation "P_n holds asymptotically almost surely (a.a.s.)" would be appropriate. We will use the notation "almost surely" (a.s.), which is standard in the random graph literature [22].

[3]In [1] the term *giant component* is defined slightly differently.

THEOREM 5.5 (see [179]). *In the random graph $\mathcal{Q}^n_{\alpha,\lambda}$ the probability $\lambda^* = 1 - \sqrt[\alpha-1]{\alpha^{-1}}$ is the threshold value for connectivity. That is, a.s. no random graph is connected for $\lambda < \lambda^*$ and a.s. every random graph is connected for $\lambda > \lambda^*$.*

The above theorem is in fact well known for binary n-cubes. However, the proof in [22] is based on an estimation of some edge boundary and utilizes in this context an isoperimetric inequality from [102].

The proof of Theorem 5.5 does in fact explicitly construct "many" independent paths which eventually lead to the desired result: Let P, Q be arbitrary vertices of the random graph. As in Theorem 5.4 we can reduce the case to P, Q having finite Hamming distance. For $\lambda > \lambda^*$, one then shows that any vertex has an arbitrary finite number of neighbors in the random graph. Using these neighboring vertices one proceeds in a manner completely analogous to the proof of Theorem 5.4. To prove that λ^* is a threshold value we show that there exist isolated vertices in the case of $\lambda < \lambda^*$. This can be proved by considering the random variable counting the isolated vertices, Z. It is obvious that Z has mean $\mu = \lambda\, \alpha^n\, (1-\lambda)^{(\alpha-1)n}$ and for finite μ one can show that Z becomes Poisson in the limit of large n. From this we can conclude that for $\lambda < \lambda^*$ and arbitrary natural number ℓ, there are at least ℓ isolated vertices in the random graph a.s.

Extensive computational studies on RNA landscapes indeed show that the neutrality $\nu/(\alpha-1)n$ is above the threshold value λ^* for many RNA structures and that there are indeed extensive (almost) connected neutral networks [95, 96].

5.2.4. Neutral Paths. Neutral paths were used to gain information about the structure of the (connected components of) neutral networks in a series of computer experiments on RNA folding landscapes [95, 96, 200]. In each step we attempt to find a neutral neighbor such that the distance from the starting point increases.

In a random subgraph $\Gamma_{n,p}$ of a distance regular graph[4] Γ_n the probability that a neutral path with d steps cannot be elongated any further equals $(1-p)^{\alpha(d)}$, where $\alpha(d)$ denotes the number of "forward steps," i.e., the number of adjacent vertices actually increasing the distance to the starting point. For Hamming graphs $\mathcal{Q}^n_\alpha$, for instance, we have $\alpha(d) = (a-1)(n-d)$. The probability that a neutral path of $\Gamma_{n,p}$ terminates after exactly d steps is given by (see [188])

$$\text{Prob}[\mathcal{L} = d] = (1-p)^{\alpha(d)} \times \prod_{d'=1}^{d} \left[1 - (1-p)^{\alpha(d'-1)}\right]. \tag{5.3}$$

Next, let $L(n) = \text{diam}(\Gamma_n)$ and $x = d/L(n)$. We consider a sequence Γ_n of distance transitive graphs such that $L = L(n) \to \infty$ and $\alpha(d)$ can be written in the form $\alpha(d) = D(n)\vartheta(x)$, where $D(n)$ is the vertex degree of Γ_n and $\vartheta : [0,1] \to [0,1]$ is differentiable. Let ω_n tend to infinity arbitrarily slowly and suppose $\zeta = \lim_{n\to\infty} \log D(n)/\log L(n)$ exists. Then $\text{Prob}[\mathcal{L}/\text{diam}(\Gamma_n) = x]$ has a maximum at

$$x^* = \begin{cases} 1 & \text{if} \quad p = \omega_n \log D(n)/D(n), \\ 0 < s < 1 & \text{if} \quad p = C \log D(n)/D(n), \quad C > 1/\zeta, \\ 0 & \text{if} \quad p = C \log D(n)/D(n), \quad C < 1/\zeta. \end{cases} \tag{5.4}$$

For an example, see Figure 5.3. Neutral paths provided the first evidence for extended neutral networks in RNA models [200].

[4] A connected graph is *distance transitive* if its automorphism group acts transitively on $\mathcal{D}_i = \{(x,y) \mid d(x,y) = i\}$ for $i \le \text{diam}(\Gamma_n)$. It is *distance regular* if the distance classes $\mathcal{D}_i$ form a coherent configuration. Distance transitivity implies distance regularity.

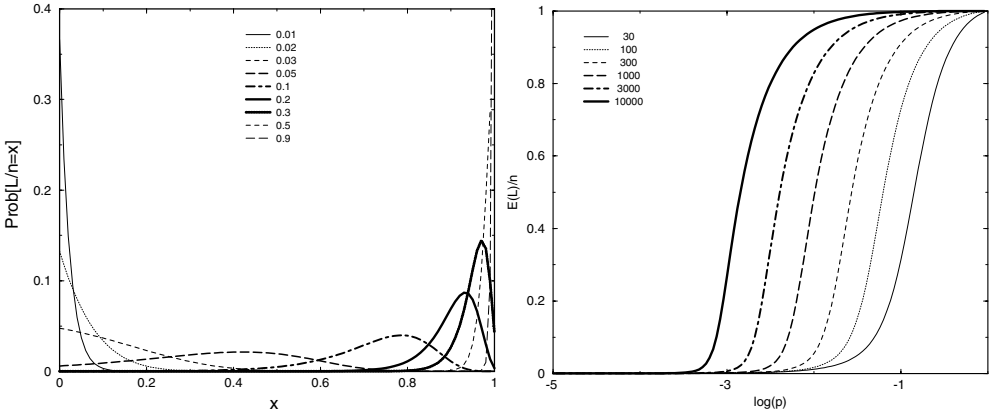
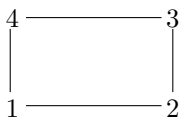


Fig. 5.3 *(left) Distribution of a scaled neutral path lengths L/n on random subgraphs of $\mathcal{Q}_2^{100}$ for different values of p. (right) Average length of a neutral path for different sequence length n.*

5.3. Sequential Dynamical Systems.

5.3.1. Genotypes and Phenotypes of SDSs. The schedules are the SDS-genotypes and form the update graph $U(Y)$ whose vertices are permutations, written as n-tuples without repetition: $(i_1, \ldots, i_n)$. In order to understand what adjacency of schedules means, let us consider $Y = \mathrm{Circ}_4$:



Now suppose we apply the maps $F_1, \ldots, F_4$ according to the orderings $(1, 2, 3, 4)$ and $(3, 2, 1, 4)$. Since 2 is adjacent to 1 and 3 there exist maps $F_1, \ldots, F_4$ such that $F_1 \circ F_2 \circ F_3 \circ F_4 \neq F_3 \circ F_2 \circ F_1 \circ F_4$ (see Figure 5.4 for an illustration). That is, in general we cannot transpose vertices 1 and 3 although they are *not* adjacent in Circ_4. Hence, if we want to define an adjacency relation between two permutations $\pi = (i_1, \ldots, i_n)$, $\pi' = (j_1, \ldots, j_n)$ such that $\prod_{r=1}^{n} F_{\pi(r),Y} = \prod_{r=1}^{n} F_{\pi'(r),Y}$ holds for all maps $F_1, \ldots, F_n$, we can *at most* allow the transposition of *consecutive* coordinates (being Y-vertices) i_k, i_{k+1} in $(i_1, \ldots, i_n)$. In fact it is straightforward to show that every transposition of consecutive nonadjacent vertices i_k, i_{k+1} leaves the SDS invariant [16]. Hence two schedules $(i_1, \ldots, i_n)$ and $(h_1, \ldots, h_n)$ are adjacent (point mutants) if they differ by exactly one flip of two consecutive coordinates that are not Y-neighbors (or equivalently if and only if (a) $i_\ell = h_\ell$, $\ell \neq k, k+1$ and (b) $\{i_k, i_{k+1}\}$ is not adjacent in Y). Note that the above definition of adjacency leads to a maximum of $n-1$ neighbors and a maximal distance of $\binom{n}{2}$ between two schedules.

$U(Y)$ induces equivalence classes of schedules by identifying any two vertices that are connected by a path in the update graph, which we will write as $\pi \sim_Y \pi'$. For an illustration, see Figure 5.4, where we draw the update graph of the square, $Y = \mathrm{Circ}_4$. It turns out that there exists a one-to-one correspondence between sets of equivalent schedules and the acyclic orientations of Y, $\psi_Y : \mathsf{S}_n / \sim_Y \longrightarrow \mathrm{Acyc}(Y)$; see Figure 5.5 [180]. An acyclic orientation of an undirected graph Y is obtained by assigning directions on its edges such that the resulting directed graph is a tree.

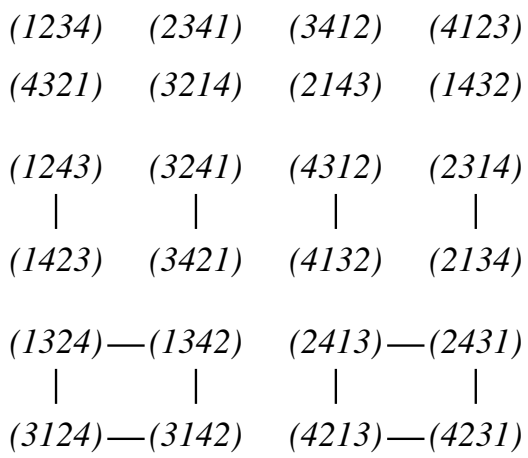


Fig. 5.4 *The update graph of $Circ_4$, $U(Circ_4)$. $U(Circ_4)$ has 24 vertices, eight of which are isolated points (corresponding to the Hamiltonian paths in $Circ_4$), four components of size 2, and two components of size 4.*

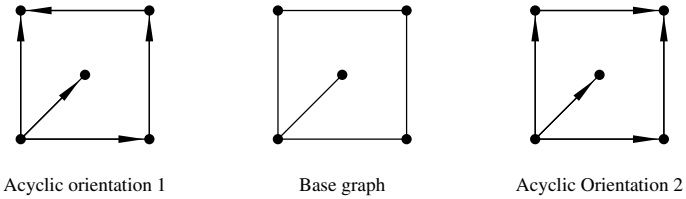


Fig. 5.5 *Acyclic orientations of an undirected graph are obtained by assigning directions to its edges such that the resulting directed graph is cycle-free.*

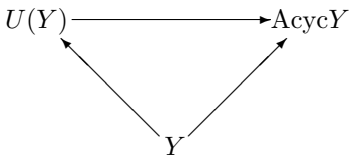
We denote the set of acyclic orientations of Y by $\mathrm{Acyc}(Y)$. The number of acyclic orientations of a graph Y is given by the absolute value of its chromatic polynomial at (-1) [222]. Accordingly, the acyclic orientations of Y can be viewed as the *phenotypes* of the mapping

$$\lambda_Y : U(Y) \longrightarrow \mathrm{Acyc}(Y). \tag{5.5}$$

It is worth mentioning that λ_Y is highly nontrivial. For example, determining the λ_Y-preimage of an acyclic orientation corresponds to the computation of the number of linear orders that are compatible to a given partial order (which is naturally induced by the acyclic orientation).

Further, it is clear that SDSs *by construction* allow for a variety of genotype-phenotype maps. Unlike the RNA case, where we have fixed genotypes and there are only relatively few concepts of phenotypes, as with, for example, secondary structures or tertiary structures, SDS genotype-phenotype maps depend on the choice of the base graph, Y. From this point of view RNA seems to be a particular case since of course the secondary structure notion strongly depends on the linear sequence realizing it.

For SDSs the graph Y yields both the update graph $U(Y)$ whose vertices are the genotypes and the phenotypes that factor through acyclic orientations:



5.3.2. Neutral networks. Now we can ask how well we can search for a specific schedule of a simulation by simply using some analogue of point mutations in the update graph. Of course, this question is motivated by our findings in the RNA case. In fact, straightforward comparison of (generalized) n-cubes $\mathcal{Q}_\alpha^n$ (the search space for RNA) and the update graphs (the search space for SDSs) reveals that both are excellent search spaces. Explicitly we have $\deg(\mathcal{Q}_\alpha^n) = (\alpha-1)n$, $\deg(U(Y)) \leq n-1$, $\mathrm{diam}(\mathcal{Q}_\alpha^n) = n$, and $\mathrm{diam}(U(Y)) \leq \binom{n}{2}$. Furthermore $|\mathcal{Q}_\alpha^n| = \alpha^n$ and $|U(Y)| = n!$ hold. In view of this similarity, it remains to analyze whether the genotype-phenotype mapping constructed in (5.5) exhibits a significant degree of neutrality.

We will next give a concentration result for the number of acyclic orientations of the random graph $G_{n,p}$ from which we can draw the following two conclusions:

(a) the number of neutral networks is sharply concentrated around its mean;
(b) the average size of a neutral network N is given by

$$p^{-n}\prod_{i=1}^{n}[1-(1-p)^i]. \tag{5.6}$$

THEOREM 5.6 (see [180]). *Let $G_{n,p}$ be a random graph, i.e., the graph over $\{1,\ldots,n\}$ where each edge is selected with independent probability p, and let $\log_2(|\mathrm{Acyc}(G_{n,p})|) : G_{n,p} \to \mathbb{N}$ be the random variable counting the number of acyclic orientations of $G_{n,p}$. Then $\log_2(|\mathrm{Acyc}(G_{n,p})|)$ is sharply concentrated around its mean, i.e., for all $\lambda > 0$,*

$$\mu_{n,p}(\{\ |\log_2(|\mathrm{Acyc}(G_{n,p})|) - \mathbb{E}[\log_2(|\mathrm{Acyc}(G_{n,p})|)]\ | > \lambda\sqrt{n(n-1)/2}\ \}) < 2e^{-\lambda^2/2},$$

where $n[\log_2(n) - \log_2 e - \log_2 p - o(1)] \leq \mathbb{E}[\log_2(|\mathrm{Acyc}(G_{n,p})|)]$. In particular on the average there are $p^{-n}\prod_{i=1}^{n}[1-(1-p)^i]$ permutations (schedules) that are mapped by λ_Y into an acyclic orientation.

Note that the above theorem, however, does not provide information on the distribution of sizes of neutral networks.

6. Dynamics on Landscapes.

6.1. Landscape Structure and Simulated Annealing. Simulated annealing [129, 98, 171] is a very general optimization method based on stochastically simulating the slow cooling of a physical system. The basic idea is that there is a "temperature" T, various ways to change the state of the system, and a probability of accepting a change that depends on the difference in the fitness function. The transition matrix is therefore of the form

$$\mathbf{P}_{yx} = \mathbf{T}_{yx} \times \begin{cases} 1 & \text{if} \quad f(y) \leq f(x), \\ \exp\big(-(f(y)-f(x))/T\big) & \text{if} \quad f(y) > f(x). \end{cases} \tag{6.1}$$

The temperature is slowly decreased at each step. The sequence $\{T_t\}$ is called the *cooling schedule*.

When the temperature is zero, changes are accepted only if f decreases, an algorithm also known as hill-climbing, adaptive walk [85], or, more generally, the greedy algorithm or steepest descent. In this case the system soon reaches a state in which none of the proposed changes can decrease the cost function, but this is usually a poor optimum. Little is known about the relationships of adaptive and gradient walks and landscape structure apart from extensive numerical studies, mostly on Nk model landscapes [68, 125, 244] and uncorrelated random landscapes [147, 146, 175]. Similar numerical studies have been performed for RNA folding landscapes [73] and in a model of early vascular land plants [166].

Landscape characteristics such as depth and difficulty determine the asymptotic behavior of simulated annealing.

THEOREM 6.1 (see [98]). *Simulated annealing converges almost surely to a global minimum if and only if the cooling schedule T_k satisfies $\sum_{k\geq 0} \exp(-\mathsf{D}/T_k) = \infty$.*

A general theory of "simulated annealing algorithms and Markov chains with rare transitions" that emphasizes the importance of depth and difficulty for convergence results and error bounds is reviewed in [31].

6.2. Quasi-Species Dynamics. A particular class of dynamics that has been studied in various landscapes is a subclass of genetic algorithms in which only mutation but no crossover is considered. Given some landscape $f : X \to \mathbb{R}$ some configuration x is replicated with rate $f(x)$. The replication process is error-prone and produces the mutant configuration y with probability $Q_{x,y}$. We can visualize this process as follows:

$$x \xrightarrow{f(x)} \xrightarrow{Q_{x,y}} x + y \tag{6.2}$$

The first class of landscapes in which mutation-based dynamics were investigated were single-peak landscapes. In a single-peak landscape one particular configuration has the maximum fitness while all other configurations have inferior fitness values. Eigen, Schuster, and collaborators [50, 52, 51, 224] completely analyzed the error-prone replication of haploid organisms (or, equivalently, biopolymer sequences) on a single-peak landscape. They discovered the genotypic error-threshold phenomenon, i.e., the existence of some critical error rate at which the population becomes unstable and drifts essentially randomly through sequence space.

More complicated landscapes were considered, beginning with the double-peak model [202] exhibiting a trade-off between width and height of the peaks that depends on the mutation rate. The quasi-species dynamics of the symmetric geometric landscape was solved in closed form in [250]. Error-threshold phenomena on spin-glass-type landscapes were studied, for instance, in [23, 227]. The effects of population sizes were discussed in [3, 29, 148, 169]. Inspired by a series of computer simulations based on RNA folding landscapes [70, 71, 73, 118], researchers have recently shifted their focus to so-called *single shape landscapes*. These arise from neutral networks by assigning a high fitness value to all sequences belonging to a particular neutral network and a low fitness to all other configurations [186].

It could be shown that RNA mutation dynamics exhibits for single-shape landscapes *phenotypic error-thresholds* [118, 186]. The phenotypic error-threshold is a natural generalization of Eigen's genotypic error-threshold, which can be considered as a phenotypic error-threshold in a landscape where the corresponding master sequence has no neutral neighbors. Accordingly, the error-threshold phenomenon does

not seem to be an artifact of the particular choice of single-peak landscapes, although many types of fitness function do not exhibit the error-threshold phenomenon [236, 248].

In the following we use a single-shape landscape of an SDS of the form

$$f_{a_0} : \mathrm{Acyc}(Y) \to \{1, \sigma\}, \quad f(a) = \begin{cases} \sigma > 1 & \text{for } a = a_0, \\ 1 & \text{else} \end{cases} \tag{6.3}$$

as an example. We shall see that the landscape

$$\lambda_Y : U(Y) \longrightarrow \mathrm{Acyc}(Y) \to \{1, \sigma\}$$

exhibits an error-threshold phenomenon for mutation-based replication of update schedules.

To this end we introduce a replication-deletion process over permutations with acyclic orientations as phenotypes. We refer to a permutation π as a *master* or a *nonmaster*, respectively, depending on whether or not π is an element of $\lambda_Y^{-1}(a_0)$. A *population* $V = \{v_i |\ i \in \mathbb{N}_N\}$ is a finite family of vertices. Each element of V has a fitness of 1 if it corresponds to a nonmaster acyclic orientation or σ otherwise. The replication-deletion process consists of two coupled random events: an element of V is selected with some fitness weighted probability and is then subject to replication, whereas another randomly chosen one is deleted. This process is the well-known Moran model [159]. In detail, the replication-deletion process works as follows [88]: suppose there are $m \in \mathbb{N}_N$ elements having the master phenotype and let $\overline{\sigma} = (N + (\sigma - 1)m)/N$. We select an ordered pair (v_r, v_d) from V: The first element v_r then has a master phenotype with probability $p_\mu = \sigma m/(N\overline{\sigma})$ and has a nonmaster phenotype otherwise. The second element v_d is chosen with uniform probability $1/(N-1)$ from $V \setminus \{v_r\}$. The pair (v_r, v_d) is mapped into the pair (v_r, v^*), where v_r remains (unchanged) in the population and v_d is replaced by v^*. In order to describe the mapping $(v_r, v_d) \mapsto (v_r, v^*)$ we first introduce the maps

$$e_j : \mathsf{S}_n \to \mathsf{S}_n; \qquad e_j((i_1, \ldots, i_n)) = (i_1, \ldots, i_{j+1}, i_j, \ldots, i_n) \tag{6.4}$$

for all $j = 1, \ldots, n$. The maps e_j are the analogue of point mutations in the RNA case. The mapping $(v_r, v_d) \mapsto (v_r, v^*)$ is now obtained as follows: We select each e_j, $j = 1, \ldots, n-1$, independently with probability φ and derive the multiset $I = (j_1, \ldots, j_s)$, where $j_a < j_b$ for $a < b$ and $j_h \in \{1, \ldots, n-1\}$. Then we set

$$v^* = \left[\prod_{j \in I} e_j \right] (v_r). \tag{6.5}$$

The above mappings are considered as independent events, and the time interval Δt which elapses between two such actions is assumed to be exponentially distributed according to $P(\Delta t > \tau) = \exp(-\tau N \overline{\sigma})$.

Let us now turn to the time evolution of a population of permutations. We introduce the following metric on acyclic orientations:

$$d(\ ,\ ) : \mathrm{Acyc}(Y) \times \mathrm{Acyc}(Y) \to \mathbb{N},$$
$$d(a, a') = |\{y \in \mathrm{E}[Y] \mid y \text{ has a different orientation in } a \text{ and } a'\}|.$$

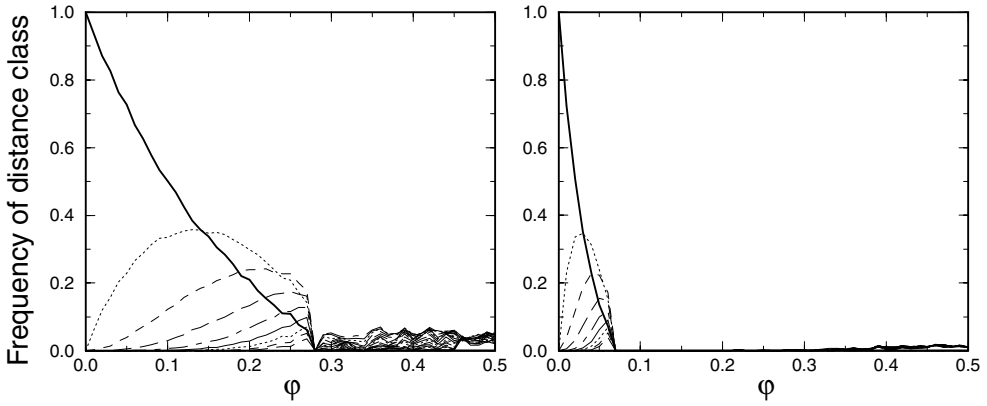


Fig. 6.1 *The error-threshold phenomenon for the compositum $f_{a_0} \circ \lambda_{G_{30,p}}$ for probabilities $p = 0.25$ (left) and $p = 1.0$ (right). We plot the fraction of a population of size $N = 1000$ in the distance classes $D = 0$ (full line), $D = 1$ (dotted), $D = 2$ (short-dashed), $D = 3$ (long-dashed), $D = 4$ (dash-dotted), etc., as a function of the mutation probability φ.*

Each element of $v \in V$ has distance $d(v, a_0)$ from the "target" a_0. In the following we focus on the frequency distribution of these distances. Specifically, we generate a random graph $Y \in G_{n,p}$ and choose a random acyclic orientation $a_0 \in \mathrm{Acyc}(Y)$ to be the master phenotype and set $\sigma = 10$. The initial population consists of $N = 1000$ permutations corresponding to the master acyclic orientation; see Figure 6.1 for two examples.

The critical value of the mutation probability φ can be obtained analytically [135]:

$$\varphi^* = 1/p(1 - \sqrt[n-1]{1/\sigma}). \tag{6.6}$$

For error rates below the threshold we have a nonuniform distribution of the population, a quasi species of schedules, and above the threshold the population is uniformly distributed. In particular, for $Y = K_n$, i.e., $p = 1$, the mapping λ_Y is bijective and, consequently, each phenotype is represented by exactly one genotype. Accordingly, we obtain the classical single-peak landscape of [51]. By modulating the edge-picking probability p one obtains a variety of genotype-phenotype mappings. The smaller p becomes, the more selective is the neutrality that is exhibited which allows to tolerate more and more replication errors.

Neutrality has a number of important impacts on the dynamical behavior of population replicating according to (6.2). The diffusive motion of the population's "center of gravity" is described in [46]. The diffusion constant is related to population size N per digit mutation rate φ and the fraction p of neutral neighbors [118], and is

$$D \approx \frac{6f(a_0)\varphi}{3 + 4N\varphi} np \tag{6.7}$$

in the case of a 4-letter alphabet. A constant "rate of innovation" was reported in [117] for the landscapes in which all neutral networks come close together, as in the case of RNA. An analytical study of very simple model landscapes shows that crossing entropy barriers can be faster by orders of magnitude than the crossing of fitness barrier [231, 249].

6.3. Genetic Algorithms and Genetic Programming. Genetic algorithms, evolutionary strategies, and genetic programming [20, 69, 111, 124, 136, 178] can be viewed as dynamical systems defined on a fitness landscape, and the interplay of landscape structure and performance of genetic algorithms is an area of active research. Most of the literature on this topic, however, deals with computer simulations and empirical connections between measures such as fitness distance correlation [120] and algorithm performance. We deliberately exclude this topic here and refer the reader to recent books, including [11, 91, 156, 205].

Much of the mathematical analysis of genetic algorithms is concerned with the convergence of the population, e.g., [167, 194]. Schemata, i.e., hyperplanes in $\mathcal{Q}_a^n$, appear to play an important role here [61, 89, 90, 92, 142, 172, 234, 235].

The fitness function $f : X \to \mathbb{R}$ can be extended in a natural way to arbitrary subsets of X by setting

$$f(A) = \frac{1}{|A|} \sum_{x \in A} f(x). \tag{6.8}$$

A schema is defined in terms of its fixed bits h as

$$\mathcal{H} = H[h] = \{x \in X | \forall i \in H \mid x_i = h_i\}. \tag{6.9}$$

Note that we regard $H \subseteq \{1, \ldots, n\}$ as the index set of fixed positions. The value $f(H[h])$ is called the schema-fitness. For a discussion of the schema theorem and the building block hypothesis we refer to the literature [2, 20, 77, 110, 111, 223]. A variety of landscape classes can be defined in terms of schema-fitnesses. We restrict ourselves to a simple example here just to give the flavor. In a deceptive landscape an optimal schema of some size is "contradicted" by one of its subschemata. Intuitively, this is just the converse of GA-easy [141].

DEFINITION 6.2 (see [247]). *A landscape f is* deceptive *if there are vertices $x, y \in \mathcal{Q}_\alpha^n$ and index sets $H \subset K \subset \{1, \ldots, n\}$ such that*

(i) $K[x] \neq K[y]$,
(ii) $f(H[x]) > f(H[z])$ *for all* z *with* $H[x] \neq H[z]$, *and*
(iii) $f(K[y]) > f(K[z])$ *for all* z *with* $K[y] \neq K[z]$.

A discussion of various notions of deceptive and GA-easy functions and their mutual relationships can be found in [162, 214].

7. Trends in Landscape Theory. In the following we will try to discuss some developments in landscape theory, currently under investigation, which we think have some relevance for a more complete picture of this subject. As this section is intended to be an outlook on different directions of landscape theory, our presentation is not entirely self-contained. Explicitly, section 7.3 does require some background on basic cohomology theory [155].

7.1. Configuration Space Topologies. Combinatorial ("discrete") landscapes are treated quite differently from their manifold ("continuous") counterparts. The reason is that functions on $\mathbb{R}^n$, or more generally Riemannian manifolds, can be analyzed in terms of differential operators such as gradients, while finite sets are usually discussed in terms of graph-theoretical properties. It seems desirable, therefore, to find a basic framework that allows us to deal with landscapes on arbitrary configuration spaces. A suitably general language is provided by the theory of *pretopological spaces*.

A pretopological space consists of an arbitrary set X and a collection $\mathcal{N}(x)$ of neighborhoods for every point $x \in X$, such that

(P1) $N \in \mathcal{N}(x)$ implies $x \in N$;
(P2) $N \in \mathcal{N}(x)$ and $N \subseteq N'$ implies $N' \in \mathcal{N}(x)$;
(P2) $N, N' \in \mathcal{N}(x)$ implies $N \cap N' \in \mathcal{N}(x)$.
Pretopologies are more general than the much more familiar topological spaces. In fact, $(X, \mathcal{N})$ is a topological space if and only if

(T) for each $N \in \mathcal{N}(x)$ there is an $N' \in \mathcal{N}(x)$ such that $N \in \mathcal{N}(y)$ for all $y \in N'$.

Directed graphs are exactly the finite pretopological spaces. Their neighborhood systems consist of all sets N' containing x and all vertices adjacent to x, i.e., $N(x) \cup \{x\} \subseteq N'$. Notions such as minima, maxima, or continuity of a function, connectedness, convergence, limits, etc., can be defined on pretopological spaces [32, 63, 126, 127]. Their usefulness in the context of genotype-phenotype maps and fitness landscapes is discussed in [208]. A further generalization to generalized closure spaces [42, 87, 99] is necessary when recombination is considered. This approach is explored in [207].

The virtue of the topological approach is that it allows a unified description of combinatorial landscapes and potential surfaces on manifolds alike. For example, basic notions such as local minima, saddle points, and so on can be defined in the same way. For example, $x \in X$ is a *local minimum* if there is a neighborhood $N \in \mathcal{N}(x)$ of x such that $f(x) \leq f(y)$ for all $y \in N$. Similarly, one can use paths to define saddle points analogous to (4.8).

Discrete Morse theory as studied in [74, 75, 76] has potentially some relevance to the landscape theory, in particular in the context of approximating landscapes. So far, however, this connection has not been explored systematically.

7.2. Modularity. The discussion of landscape and genotype-phenotype maps so far has made little use of the internal structure of the underlying configurations. In some cases, however, an optimization problem can be simplified by the observation that it suffices to optimize partial configurations separately. A trivial example is the additive fitness landscape $f(x) = \sum_i a_i x_i$ on the binary hypercube, where the global optimum can be found by maximizing $a_i x_i$ for each i. In biology one observes that phenotypes are usually composed of "modules" that can evolve approximately independently. The notion of modularity also plays an important role in evolutionary programming. For a very recent overview we refer to the book [28]. Given the practical importance of this topic it is perhaps surprising that there is no generally applicable definition of modularity.

A first attempt by Frenken, Marengo, and Valente [79] is restricted to sequence spaces $X = \mathcal{Q}_\alpha^n$ at present. Given a schema $\mathcal{H} = H[h]$ with fixed positions H and fixed bits $h \subseteq \{1, \ldots, n\}$ and a sequence x, the *projection* of x onto $\mathcal{H}$ is the sequence with positions

$$(x \wedge \mathcal{H})_i = \begin{cases} h_i & \text{if} \quad i \in H, \\ x_i & \text{if} \quad i \notin H. \end{cases} \tag{7.1}$$

A schema $\mathcal{H}$ is *dominant* for a fitness function $f : X \to \mathbb{R}$ if $f(x \wedge \mathcal{H}) \geq f(x)$ for all $x \in X$. A *cover* of f is a collection $\mathfrak{H}$ of schemata such that (i) each schema $\mathcal{H} = H[h] \in \mathfrak{H}$ is dominant and (ii) for each $i \in \{1, \ldots, n\}$ there is $\mathcal{H} = H[h] \in \mathfrak{H}$ with $i \in H$. If a cover is known, the global optimum can easily be obtained by a series of projections. A *module* can now be defined as a set of positions $M \subset \{1, \ldots, n\}$ such that there is a cover $\mathfrak{H}$ for f with the property that either $H \subseteq M$ or $H \cap M = \emptyset$ for all dominating schemata $H[h] \in \mathfrak{H}$.

A more general approach to this topic naturally starts by extending the set X of configurations by their "components" (subsequences, subtours, etc.) to a partially ordered set $(\mathbb{X}, \preceq)$ such that $y \preceq x$ means "y is a part of x." This lattice has configurations $x \in X$ as its maximal elements. Landscape decompositions of the form

$$f(x) = \sum_{y \preceq x} \psi(y) \tag{7.2}$$

are then of immediate interest. In the case of sequence spaces a component y may be identified with a schema, i.e., a set of fixed positions and their fixed bit values. The fitness function f is extended from X to $\mathbb{X}$ by means of (6.8). The function $\psi : \mathbb{X} \to \mathbb{R}$ is then given by the Möbius inversion formula

$$\psi(x) = \sum_{y \preceq x} \mu(x, y) f(y), \tag{7.3}$$

where the *Möbius function* is obtained from the recursion [193]

$$\mu(x, y) = - \sum_{x \preceq z \preceq z} \mu(x, z) \qquad \text{with} \quad \mu(x, x) = 1. \tag{7.4}$$

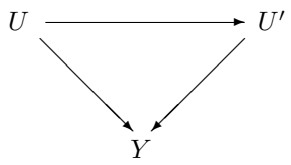
In this particular case one can interpret the values of $\psi(y)$ as the "residual schema fitnesses."

Since f is originally defined on X only, it can be extended to $\mathbb{X}$ also in other ways. For instance, one might require that the functions ψ are nonzero only on *small* components.

7.3. An Algebraic Context for Neutral Landscapes. From an algebraic point of view it is natural to ask a question such as the following: *Given a graph Y, under which conditions is there a specific* class *of landscapes f on Y, and if so, how many landscapes of this* class *can be constructed?* Typically one would be interested in Y-local properties such as having a fixed number k of neutral Y-neighbors; see Figure 7.1 for an example. Surprisingly, it appears that no such theory has been developed for landscapes yet. Since we will frequently make explicit references to the underlying graph in the following, we shall write $\mathrm{v}[Y]$ and $\mathrm{e}[Y]$ for its vertex and edge-sets, respectively.

For some graph Y let $B_1(j)$ denote the set of j neighbors in Y. Call a landscape k-neutral over Y if for any $j \in \mathrm{v}[Y]$ we have that $|\{\, i \in B_1(j) \mid f(i) = f(j) \,\}| = k$ holds. We are interested in the collection of all k-neutral landscapes over Y. The key idea is to consider the class of *all* induced subgraphs of Y and relate the k-neutral landscapes over these to the k-neutral landscapes over Y. More precisely, we expand k-neutral landscapes over Y-subgraphs to k-neutral landscapes over Y. Clearly, this idea is motivated from analytic continuations of functions.

We begin our analysis by introducing what we consider to be the "local pieces" of Y: To this end we consider the category $C(Y)$, whose objects are all Y-induced subgraphs and whose morphisms are the inclusion maps. In other words, we have the commutative diagram



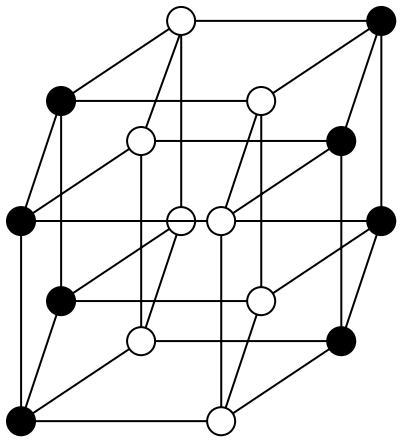


Fig. 7.1 *A 2-neutral landscape $f : \mathcal{Q}_2^4 \to \mathbb{F}_2$. Here all black vertices and all white vertices map into 0 and 1, respectively. Note that each of the two fitness classes forms two connected subgraphs.*

Let $X \longrightarrow Y$ be an object in $C(Y)$. Then a covering of X is a multiset of $C(Y)$-morphisms $(U_i \longrightarrow X)_{i\in I}$ with the property $\cup_I \mathrm{e}[U_i] = \mathrm{e}[X]$. The class of all such coverings of all such objects is called the I-topology[5] on $C(Y)$. Next we consider the mapping P_k, which assigns to each induced Y-subgraph the free Abelian group generated by its k-neutral maps:

$$P_k(U) = \mathbb{Z}[\{f : Y \to K \mid \forall\, j \in U;\ f \text{ is } k\text{-neutral in j }\}]. \tag{7.5}$$

We will denote elements of $P(U)$ by f_U. For any $U, U' \in C(Y)$ let $\mathrm{res}_U(f_{U'}) = f_{U\cap U'} \in P(U \cap U')$ be the mapping $f_{U'} \mapsto f_{U'}$, considered as an element of $P(U \cap U')$, which is naturally induced by the inclusion $U \cap U' \to U'$. Accordingly, any $C(Y)$-morphism induces the commutative diagram

$$\begin{array}{ccc} U & \longrightarrow & P_k(U) \\ \downarrow & & \uparrow \mathrm{res}_U \\ U' & \longrightarrow & P_k(U') \end{array}$$

and P_k is a contravariant functor from $C(Y)$ into the Abelian groups **Ab**. Next we introduce the derivation map[6]

$$d^{(s)} : \prod_{\cap_{h=1}^{s} U_{i_h}, i_h < i_{h+1}} P_k(\cap_{h=1}^{s} U_{i_h}) \longrightarrow \prod_{\cap_{h=1}^{s+1} U_{i_h}, i_h < i_{h+1}} P_k(\cap_{h=1}^{s+1} U_{i_h}),$$

where $d^{(s)}$ is defined as follows: $d^{(s)}(f_{\cap_{h=1}^{s} U_{i_h}}) = (\beta_{\cap_{h=1}^{s+1} U_{i_h}})$ and

$$\beta_{\cap_{h=1}^{s+1} U_{i_h}} = \sum_{g=1}^{s+1} (-1)^g \mathrm{res}_{U_{i_g}} f_{\cap_{h,h\neq g} U_{i_h}}.$$

[5]Thus named since it is induced by *inclusion maps*.

[6]This is exactly the derivation map for the standard Cech cohomology. Note, however, that depending on the nature of the base category (which could contain morphisms that are not inclusion maps) the notion of res in what follows can become ambiguous, as these maps are no longer restriction maps.

The key result relating local and global information on k-neutral landscapes can be encoded in the exactness of the following short sequence of Abelian groups.

LEMMA 7.1 (see [181]). *Let Y be a connected graph and $(U_i \longrightarrow Y)_{i\in I}$ be a covering of Y. Then we have the short exact sequence*

$$P_k(Y) \longrightarrow^{d^{(1)}} \prod_{i\in I} P_k(U_i) \longrightarrow^{d^{(2)}} \prod_{i<j} P_k(U_i \cap U_j). \tag{7.6}$$

Equivalently, the contravariant functor $P_k : C(Y) \longrightarrow \mathrm{Ab}$ is a sheaf.

Basically, Lemma 7.1 allows us to arbitrarily patch together consistent, local k-neutral pieces in order to obtain a k-neutral landscape over Y. One natural algebraic invariant of the sheaf P is its cohomology, and since P encodes information on a collection of Y-local k-neutral landscapes its cohomology groups will contain information on the neutral landscapes over Y itself. The following result provides a purely algebraic interpretation of the k-neutral landscapes over Y as a particular cohomology group of the sheaf P_k.

THEOREM 7.2 (see [181]). *Let Y be a graph. Then we have*

$$\mathbb{Z}[Neut_k(Y)] \cong H^0(Y, P_k), \tag{7.7}$$

where $H^0(Y, P_k)$ is the 0th derived functor cohomology group of the sheaf P_k.

7.4. Landscape Morphisms. Landscape theory so far is not a *relative* theory in the sense that we would understand how structural changes in the move set (or, equivalently, base graph over the configurations) affect its properties. Relative theories, however, are standard in mathematics. For instance, in algebraic geometry one has the concept of base changes of schemes over sites or considers the mappings of (co)homology groups of topological spaces. In the following we will discuss briefly the particular case of morphisms between SDSs as introduced in [184]. Of course, we will try to design SDS-morphisms such that they allow for some "information" transfer from one SDS (viewed as a dynamical system) to another.

DEFINITION 7.3. *Let $[\mathfrak{F}_Y, \pi]$ and $[\mathfrak{F}_Z, \sigma]$ be two SDSs. A morphism $\Phi : [\mathfrak{F}_Y, \pi] \longrightarrow [\mathfrak{F}_Z, \sigma]$ is a tuple (φ, ψ), where $\varphi : Y \longrightarrow Z$ and $\psi : \mathbb{G}[\mathfrak{F}_Z, \sigma] \longrightarrow \mathbb{G}[\mathfrak{F}_Y, \pi]$ are graph and digraph morphisms, respectively.*

Here $\mathbb{G}[\mathfrak{F}_Y, \sigma]$ and $\mathbb{G}[\mathfrak{F}_Y, \pi]$ denote the phase spaces of the corresponding SDSs.

Next we show that there are in fact nontrivial SDS-morphisms and study one particular class which is naturally induced by locally bijective and locally surjective graph morphisms. Here, we call a graph morphism $\varphi : Y \longrightarrow Z$ locally surjective and locally bijective, respectively, if and only if

$$\mathrm{res}_{B_{1,Y}(i)}(\varphi) : B_{1,Y}(i) \longrightarrow B_{1,Z}(\varphi(i)) \tag{7.8}$$

is surjective and bijective for all $i \in \mathrm{v}[Y]$.

Let $\varphi : Y \longrightarrow Z$ be a graph morphism. Then we call the set

$$\mathrm{Acyc}^{\varphi}(Y) = \{\mathfrak{O} \in \mathrm{Acyc}(Y) \mid \forall\, z \in \mathrm{e}[Z];\ \forall\, y, y' \in \varphi^{-1}(z);\ \mathfrak{O}_Y(y) = \mathfrak{O}_Y(y')\}$$

the set of φ-symmetric acyclic orientations. It is shown in [184] that there exists a one-to-one correspondence $\psi_\varphi : \mathrm{Acyc}(Z) \longrightarrow \mathrm{Acyc}^{\varphi}(Y)$ for locally surjective φ. From this we obtain a mapping $\eta_\varphi : \mathsf{S}_m \to \mathsf{S}_n$ such that

$$\begin{array}{ccc} \mathsf{S}_m & \xrightarrow{\eta_\varphi} & \mathsf{S}_n \\ \downarrow & & \downarrow \\ \mathrm{Acyc}(Z) & \xrightarrow{\psi'_\varphi} & \mathrm{Acyc}(Y) \end{array}$$

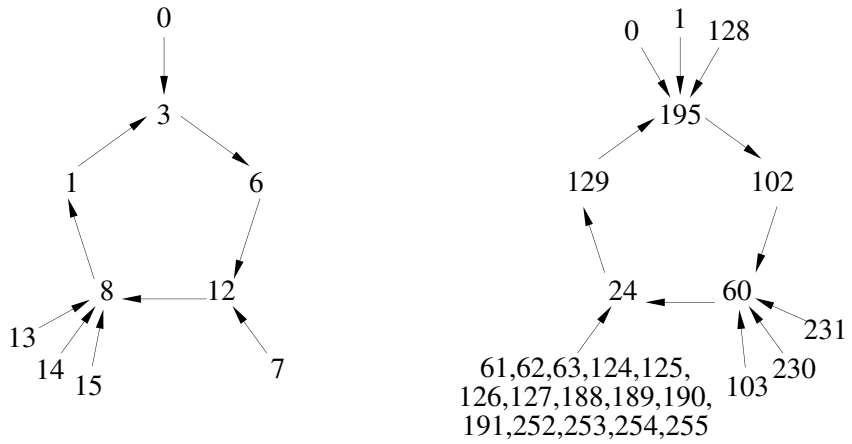


Fig. 7.2 *Illustration of Theorem* 7.4 *showing a phase space embedding induced by the covering map* $\mathcal{Q}_2^3 \to K_4$. *The figure on the left shows a limit cycle of an SDS composed by Min-functions (i.e., local, Boolean maps returning the minority value in the corresponding neighborhood and* 0 *in case of an equal number of* 0*'s and* 1*'s) over* K_4 *with the identity as update schedule. The figure on the right shows the corresponding SDS over* $\mathcal{Q}_2^3$ *with the schedule* $\eta_\varphi(id)$, *as defined in Theorem* 7.4. *It follows that the digraph on the left can be embedded into the digraph on the right.*

is commutative.[7] The following theorem establishes a relation between the phase spaces of SDS that have SDS-morphisms induced by locally bijective and surjective graph morphisms, respectively. Let $[\mathrm{Nor}_Z, \pi]$ denote an SDS over the graph Z, with a Boolean Nor function on each vertex.

THEOREM 7.4 (see [184]). *Let* Y, Z *be connected loop-free graphs, let* $\varphi : Y \longrightarrow Z$ *be a graph morphism, and define*

$$\varphi_* : \mathfrak{F}_2^{|Z|} \longrightarrow \mathfrak{F}_2^{|Y|} \qquad by \quad \varphi_*(x)_k = x_{\varphi(k)}.$$

Then the following assertions hold:

(a) *If* $\varphi : Y \longrightarrow Z$ *is locally bijective and* $[\mathfrak{F}_Z, \pi]$ *and* $[\mathfrak{F}_Y, \eta_\varphi(\pi)]$ *are induced by the set of local functions* $f_{(k)} : \mathbb{F}_2^k \to \mathbb{F}_2$, *then one has the following morphism of SDSs:*

$$\Phi = (\varphi, \varphi_*) : [\mathfrak{F}_Y, \eta_\varphi(\pi)] \longrightarrow [\mathfrak{F}_Z, \pi]. \tag{7.9}$$

(b) *Let* $\varphi : Y \longrightarrow Z$ *be locally surjective. Then we have the following morphism of SDSs:*

$$\Phi = (\varphi, \varphi_*) : [\mathrm{Nor}_Y, \eta_\varphi(\pi)] \longrightarrow [\mathrm{Nor}_Z, \pi]. \tag{7.10}$$

In particular, Theorem 7.4 shows how to translate any graph automorphism of Y into a phase space isomorphism of the corresponding SDS, independent of the Y-local functions used; see Figure 7.2. In other words, the symmetries of the base graph induce dynamically equivalent schedules, which proves in particular that SDSs can in fact be formulated over *unlabeled* graphs.

Theorem 7.4 describes the relationship between two SDSs over graphs that are related by a locally surjective graph morphism. A similar morphism concept would be a step toward a relative landscape theory. For example, one might try to identify the impact of certain classes of graph morphisms on the corresponding amplitude spectra.

[7] Here $\psi_{\varphi'}$ is the map induced by ψ_φ by identifying $\mathrm{Acyc}^\varphi(Y)$-elements as acyclic orientations.

Acknowledgments. We thank Henning Mortveit for providing Figure 7.2 and Alfred Karpfen for supplying the data for Figure 1.2.

REFERENCES

[1] M. Ajtai, J. Komlós, and E. Szemerédi, *Largest random component of a k-cube*, Combinatorica, 2 (1982), pp. 1–7.
[2] L. Altenberg, *The schema theorem and the Price's theorem*, in Foundations of Genetic Algorithms, Vol. 3, L. D. Whitley and M. D. Vose, eds., 1995, Morgan-Kaufmann, San Francisco, CA, pp. 23–49.
[3] D. Alves and J. F. Fontanari, *Error threshold in finite populations*, Phys. Rev. E, 57 (1998), pp. 7008–7013.
[4] E. Angel and V. Zissimopoulos, *Autocorrelation coefficient for the graph bipartitioning problem*, Theoret. Comput. Sci., 191 (1998), pp. 229–243.
[5] E. Angel and V. Zissimopoulos, *On the quality of local search for the quadratic assignment problem*, Discrete Appl. Math., 82 (1998), pp. 15–25.
[6] E. Angel and V. Zissimopoulos, *On the classification of NP-complete problems in terms of their correlation coefficient*, Discrete Appl. Math., 99 (2000), pp. 261–277.
[7] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi, *Complexity and Approximation: Combinatorial Optimization Problems and Their Approximability Properties*, Springer-Verlag, New York, 1999.
[8] R. Azencott, *Simulated Annealing*, John Wiley, New York, 1992.
[9] L. Babel, S. Baumann, M. Lüdecke, and G. Tinhofer, `STABCOL`*: Graph Isomorphism Testing Based on the Weisfeiler-Leman Algorithm*, Tech. Report TUM-M9702, TU München, Garching, Germany, 1997.
[10] L. Babel, I. V. Chuvaeva, M. Klin, and D. V. Pasechnik, *Algebraic Combinatorics in Mathematical Chemistry. Methods and Algorithms.* II. *Program Implementation of the Weisfeiler-Leman Algorithm*, Tech. Report TUM-M9701, TU München, Garching, Germany, 1995.
[11] W. Banzhaf, P. Nordin, R. E. Keller, and F. D. Francone, *Genetic Programming: An Introduction: On the Automatic Evolution of Computer Programs and Its Applications*, Morgan-Kaufmann, San Mateo, CA, 1998.
[12] L. Barnett, *Ruggedness and neutrality—the NKp family of fitness landscapes*, in Alive VI: Sixth International Conference on Artificial Life, C. Adami, R. K. Belew, H. Kitano, and C. E. Taylor, eds., MIT Press, Cambridge, MA, 1998, pp. 18–27.
[13] C. L. Barrett, H. S. Mortveit, and C. M. Reidys, *Elements of a theory of simulation* II: *Sequential dynamical systems*, Appl. Math. Comput., 107 (2000), pp. 121–136.
[14] C. L. Barrett, H. S. Mortveit, and C. M. Reidys, *Elements of a theory of simulation* III: *Sequential dynamical systems*, Appl. Math. Comput., 122 (2001), pp. 325–340.
[15] C. L. Barrett, H. S. Mortveit, and C. M. Reidys, *Elements of a theory of simulation* IV: *Sequential dynamical systems*, Appl. Math. Comput., to appear.
[16] C. L. Barrett and C. M. Reidys, *Elements of a theory of simulation* I: *Sequential CA over random graphs*, Appl. Math. Comput., 98 (1999), pp. 241–259.
[17] O. Bastert, D. Rockmore, P. F. Stadler, and G. Tinhofer, *Landscapes on spaces of trees*, Appl. Math. Comput., to appear.
[18] O. M. Becker and M. Karplus, *The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics*, J. Chem. Phys., 106 (1997), pp. 1495–1517.
[19] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, Vol. 1, Athena Scientific, Belmont, MA, 1995.
[20] A. D. Bethke, *Genetic Algorithms and Function Optimizers*, Ph.D. thesis, University of Michigan, Ann Arbor, MI, 1991.
[21] K. Binder and A. P. Young, *Spin glasses: Experimental facts, theoretical concepts, and open questions*, Rev. Mod. Phys., 58 (1986), pp. 801–976.
[22] B. Bollobás, *Random Graphs*, Academic Press, New York, 1985.
[23] S. Bonhoeffer, J. S. McCaskill, P. F. Stadler, and P. Schuster, *RNA multi-structure landscapes. A study based on temperature dependent partition functions*, Eur. Biophys. J., 22 (1993), pp. 13–24.
[24] A. J. Bray and M. A. Moore, *Metastable states in spin glasses with short-ranged interactions*, J. Phys. C, 14 (1981), pp. 1313–1327.

[25] A. E. Brouwer, A. M. Cohen, and A. Neumaier, *Distance-Regular Graphs*, Springer-Verlag, Berlin, New York, 1989.
[26] R. Bürger, *The Mathematical Theory of Selection, Recombination, and Mutation*, John Wiley, Chichester, UK, 2000.
[27] R. E. Burkard and E. Çela, *Linear assignment problems and extensions*, in Handbook of Combinatorial Optimization: Supplement Volume A, D. Du and P. M. Pardalos, eds., Kluwer Academic, Dordrecht, The Netherlands, 1999, pp. 75–149.
[28] W. Callebaut and D. Rasskin-Gutman, eds., *Modularity. Understanding the development and evolution of complex natural systems*, MIT Press, Cambridge, MA, 2002, to appear.
[29] P. R. A. Campos and J. F. Fontanari, *Finite-size scaling of the error threshold transition in finite population*, J. Phys. A Math. Gen., 32 (1999), pp. L1–L7.
[30] O. Catoni, *Rough large deviation estimates for simulated annealing: Application to exponential schedules*, Ann. Probab., 20 (1992), pp. 1109–1146.
[31] O. Catoni, *Simulated annealing algorithms and Markov chains with rate transitions*, in Seminaire de Probabilites XXXIII, J. Azema, M. Emery, M. Ledoux, and M. Yor, eds., Lecture Notes in Math. 709, Springer-Verlag, Berlin, Heidelberg, 1999, pp. 69–119.
[32] E. Čech, *Topological Spaces*, John Wiley, London, 1966.
[33] T. R. Cech, *Self-splicing RNA: Implications for evolution*, Int. Rev. Cytol., 93 (1985), pp. 3–22.
[34] T. R. Cech, *RNA as an enzyme*, Sci. Amer., 11 (1986), pp. 76–84.
[35] H. S. Chan and K. A. Dill, *Sequence space soup*, J. Chem. Phys., 95 (1991), pp. 3775–3787.
[36] F. R. K. Chung, *Spectral Graph Theory*, CBMS Reg. Conf. Ser. Math. 92, AMS, Providence RI, 1997.
[37] B. Codenotti and L. Margara, *Local Properties of Some NP-Complete Problems*, Tech. Report TR 92-021, International Computer Science Institute, Berkeley, CA, 1992.
[38] J. P. Crutchfield and P. Schuster, eds., *Evolutionary Dynamics—Exploring the Interplay of Selection, Neutrality, Accident, and Function*, Oxford University Press, Oxford, UK, 2002, to appear.
[39] J. C. Culberson, *Mutation-crossover isomorphism and the construction of discriminating functions*, Evol. Comp., 2 (1995), pp. 279–311.
[40] R. Das, J. P. Crutchfield, M. Mitchell, and J. E. Hanson, *Evolving globally synchronized cellular automata*, in Proceedings of the Sixth International Conference on Genetic Algorithms, L. J. Eshelman, ed., Morgan-Kaufmann, San Mateo, CA, 1995, pp. 336–343.
[41] E. B. Davies, G. M. L. Gladwell, J. Leydold, and P. F. Stadler, *Discrete nodal domain theorems*, Linear Algebra Appl., 336 (2001), pp. 51–60.
[42] M. M. Day, *Convergence, closure, and neighborhoods*, Duke Math. J., 11 (1944), pp. 181–199.
[43] V. M. de Oliveira, J. F. Fontanari, and P. F. Stadler, *Metastable states in high order short-range spin glasses*, J. Phys. A Math. Gen., 32 (1999), pp. 8793–8802.
[44] Y. C. de Verdière, *Multiplicités des valeurs propres Laplaciens discrete at Laplaciens continus*, Rend. Mat. Appl. (7), 13 (1993), pp. 433–460 (in French).
[45] B. Derrida and E. Gardner, *Metastable states of a spin glass chain at* 0 *temperature*, J. Physique, 47 (1986), pp. 959–965.
[46] B. Derrida and L. Peliti, *Evolution in a flat fitness landscape*, Bull. Math. Biol., 53 (1991), pp. 355–382.
[47] K. A. Dill, S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yeo, P. D. Thomas, and H. S. Chan, *Principles of protein folding: A perspective from simple exact models*, Prot. Sci., 4 (1995), pp. 561–602.
[48] J. P. Doye, M. A. Miller, and D. J. Welsh, *Evolution of the potential energy surface with size for Lennard-Jones clusters*, J. Chem. Phys., 111 (1999), pp. 8417–8429.
[49] A. M. Duval and V. Reiner, *Perron-Frobenius type results and discrete versions of nodal domain theorems*, Linear Algebra Appl., 294 (1999), pp. 259–268.
[50] M. Eigen, *Selforganization of matter and the evolution of biological macromolecules*, Die Naturwissenschaften, 10 (1971), pp. 465–523.
[51] M. Eigen, J. McCaskill, and P. Schuster, *The molecular quasispecies*, Adv. Chem. Phys., 75 (1989), pp. 149–263.
[52] M. Eigen and P. Schuster, *The Hypercycle*, Springer-Verlag, New York, Berlin, 1979.
[53] A. D. Ellington, *Aptamers achieve the desired recognition*, Cur. Biol., 4 (1994), pp. 427–429.
[54] A. D. Ellington and J. W. Szostak, *In vitro selection of RNA molecules that bind specific ligands*, Nature, 346 (1990), pp. 818–822.
[55] P. Erdős and A. Rényi, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kut. Int. Közl, 5 (1960), pp. 17–61.

[56] I. A. Faradžev, M. H. Klin, and M. E. Muzychuk, *Cellular rings and groups of automorphisms of graphs*, in Investigations in Algebraic Theory of Combinatorial Objects, I. A. Faradžev, A. A. Ivanov, M. H. Klin, and A. J. Woldar, eds., Math. Appl. (Soviet Ser.) 84, Kluwer Academic, Dordrecht, The Netherlands, 1994.
[57] J. S. Farris, *The logical basis of phylogenetic analysis*, in Advances in Cladistics, N. I. Platnick and V. A. Funk, eds., Columbia University Press, New York, 1983, pp. 1–36.
[58] J. Felsenstein, *Evolutionary trees from DNA sequences: A maximum likelihood approach*, J. Mol. Evol., 17 (1981), pp. 368–376.
[59] F. F. Ferreira, J. F. Fontanari, and P. F. Stadler, *Landscape statistics of the low autocorrelated binary string problem*, J. Phys. A, 33 (2000), pp. 8635–8647.
[60] M. Fiedler, *A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory*, Czechoslovak Math. J., 25 (1975), pp. 619–633.
[61] P. Field, *Non-binary transformations of genetic algorithms*, Complex Systems, 9 (1995), pp. 11–28.
[62] N. J. Fine, *The generalized Walsh functions*, Trans. Am. Math. Soc., 69 (1950), pp. 66–77.
[63] H. R. Fischer, *Limesräume*, Math. Ann., 137 (1959), pp. 269–303.
[64] C. Flamm, W. Fontana, I. Hofacker, and P. Schuster, *RNA folding kinetics at elementary step resolution*, RNA, 6 (2000), pp. 325–338.
[65] C. Flamm, I. L. Hofacker, S. Maurer-Stroh, P. F. Stadler, and M. Zehl, *Design of multi-stable RNA molecules*, RNA, 7 (2001), pp. 254–265.
[66] C. Flamm, I. L. Hofacker, and P. F. Stadler, *RNA in silico: The computational biology of RNA secondary structures*, Adv. Complex Systems, 2 (1999), pp. 65–90.
[67] C. Flamm, I. L. Hofacker, P. F. Stadler, and M. T. Wolfinger, *Barrier trees of degenerate landscapes*, Z. Phys. Chem., to appear.
[68] H. Flyvbjerg and B. Lautrup, *Evolution in a rugged fitness landscape*, Phys. Rev. A, 46 (1992), pp. 6714–6723.
[69] D. B. Fogel, *Evolutionary Computation*, IEEE Press, New York, 1995.
[70] W. Fontana, W. Schnabl, and P. Schuster, *Physical aspects of evolutionary optimization and adaption*, Phys. Rev. A, 40 (1989), pp. 3301–3321.
[71] W. Fontana and P. Schuster, *A computer model of evolutionary optimization*, Biophys. Chem., 26 (1987), pp. 123–147.
[72] W. Fontana and P. Schuster, *Continuity in evolution: On the nature of transitions*, Science, 280 (1998), pp. 1451–1455.
[73] W. Fontana, P. F. Stadler, E. G. Bornberg-Bauer, T. Griesmacher, I. L. Hofacker, M. Tacker, P. Tarazona, E. D. Weinberger, and P. Schuster, *RNA folding and combinatory landscapes*, Phys. Rev. E, 47 (1993), pp. 2083–2099.
[74] R. Forman, *Morse theory for cell complexes*, Adv. Math., 134 (1998), pp. 90–145.
[75] R. Forman, *Witten-Morse theory for cell complexes*, Topology, 37 (1998), pp. 945–979.
[76] R. Forman, *Morse theory and evasiveness*, Combinatorica, 20 (2000), pp. 489–504.
[77] S. Forrest and M. Mitchell, *Relative building block fitness and the building block hypothesis*, in Foundations of Genetic Algorithms, Vol. 2, L. D. Whitley, ed., Morgan-Kaufmann, San Mateo, CA, 1993, pp. 109–126.
[78] H. Frauenfelder, A. R. Bishop, A. Garcia, A. Perelson, P. Schuster, D. Sherrington, and P. J. Swart, eds., *Landscape Paradigms in Physics and Biology: Concepts, Structures, and Dynamics*, Elsevier, Amsterdam, 1997.
[79] K. Frenken, L. Marengo, and M. Valente, *Interdependecies, nearly-decomposability, and adaptation*, in Computational Techniques to Model Learning in Economics, T. Brenner, ed., Kluwer Academic, Boston, 1998.
[80] J. Friedman, *Some geometric aspects of graphs and their eigenfunctions*, Duke Math. J., 69 (1993), pp. 487–525.
[81] R. García-Pelayo and P. F. Stadler, *Correlation length, isotropy, and meta-stable states*, Phys. D, 107 (1997), pp. 240–254.
[82] M. Garey and D. Johnson, *Computers and Intractability. A Guide to the Theory of $\mathcal{NP}$ Completeness*, Freeman, San Francisco, CA, 1979.
[83] J. Garnier and L. Kallel, *Efficiency of local search with multiple local optima*, SIAM J. Discrete Math., to appear.
[84] P. Garstecki, T. X. Hoang, and M. Cieplak, *Energy landscapes, supergraphs, and "folding funnels" in spin systems*, Phys. Rev. E, 60 (1999), pp. 3219–3226.
[85] J. H. Gillespie, *Molecular evolution over the mutational landscape*, Evolution, 38 (1984), pp. 1116–1129.
[86] P. Gitchoff and G. P. Wagner, *Recombination induced hypergraphs: A new approach to mutation-recombination isomorphism*, Complexity, 2 (1996), pp. 47–43.

[87] S. Gniłka, *On extended topologies.* I: *Closure operators*, Ann. Soc. Math. Pol. Ser. I Commentat. Math., 34 (1994), pp. 81–94.
[88] U. Göbel, C. V. Forst, and P. Schuster, *Structural constraints and neutrality in RNA*, in LNCS/LNAI Proceedings of GCB96, R. Hofestädt, ed., Lecture Notes in Comput. Sci. 1278, Springer-Verlag, Berlin, Heidelberg, New York, 1997, pp. 156–165.
[89] D. E. Goldberg, *Genetic algorithms and Walsh functions. Part* I: *A gentle introduction*, Complex Systems, 3 (1989), pp. 129–152.
[90] D. E. Goldberg, *Genetic algorithms and Walsh functions. Part* II: *Deceptiveness and its analysis*, Complex Systems, 3 (1989), pp. 153–176.
[91] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA, 1989.
[92] D. E. Goldberg and M. Rudnik, *Genetic algorithms and the variance of fitness*, Complex Systems, 5 (1991), pp. 265–278.
[93] D. J. Gross and M. Mèzard, *The simplest spin glass*, Nuclear Phys. B, 240 (1984), pp. 431–452.
[94] L. K. Grover, *Local search and the local structure of NP-complete problems*, Oper. Res. Lett., 12 (1992), pp. 235–243.
[95] W. Grüner, R. Giegerich, D. Strothmann, C. M. Reidys, J. Weber, I. L. Hofacker, P. F. Stadler, and P. Schuster, *Analysis of RNA sequence structure maps by exhaustive enumeration.* I. *Neutral networks*, Monath. Chem., 127 (1996), pp. 355–374.
[96] W. Grüner, R. Giegerich, D. Strothmann, C. M. Reidys, J. Weber, I. L. Hofacker, P. F. Stadler, and P. Schuster, *Analysis of RNA sequence structure maps by exhaustive enumeration.* II. *Structures of neutral networks and shape space covering*, Monath. Chem., 127 (1996), pp. 375–389.
[97] I. Gutman, S.Klavžar, and B. Mohar, eds., *Fifty years of the Wiener index*, MATCH, 36 (1996).
[98] B. Hajek, *Cooling schedules for optimal annealing*, Math. Oper. Res., 13 (1988), pp. 311–329.
[99] P. C. Hammer, *Extended topology: Set-valued set functions*, Nieuw Arch. Wisk. (3), 10 (1962), pp. 55–77.
[100] C. Hansch, A. Leo, and D. Hoeckman, *Exploring QSAR*, American Chemical Society, Washington, DC, 1995.
[101] T. F. Hansen and G. P. Wagner, *Modelling genetic architecture: A multilinear theory of gene interaction*, Theor. Pop. Biol., 59 (2001), pp. 61–86.
[102] S. Hart, *A note on the edges on the n-cube*, Discrete Math., 14 (1976), pp. 157–163.
[103] C. Haslinger and P. F. Stadler, *RNA structures with pseudo-knots: Graph-theoretical and combinatorial properties*, Bull. Math. Biol., 61 (1999), pp. 437–467.
[104] D. Heidrich, W. Kliesch, and W. Quapp, *Properties of Chemically Interesting Potential Energy Surfaces*, Lecture Notes in Chem. 56, Springer-Verlag, Berlin, 1991.
[105] T. Hida, *Brownian Motion*, Springer-Verlag, New York, 1980.
[106] D. G. Higman, *Intersection matrices for finite permutation groups*, J. Algebra, 6 (1967), pp. 22–42.
[107] D. G. Higman, *Coherent configurations.* I: *Ordinary representation theory*, Geom. Dedicata, 4 (1975), pp. 1–32.
[108] D. G. Higman, *Coherent configurations.* II: *Weights*, Geom. Dedicata, 5 (1976), pp. 413–424.
[109] I. L. Hofacker, P. Schuster, and P. F. Stadler, *Combinatorics of RNA secondary structures*, Discrete Appl. Math., 88 (1998), pp. 207–237.
[110] J. H. Holland, *Genetic algorithms and classifier systems: Foundations and future directions*, in Proceedings of the 2nd International Conference on Genetic Algorithms, J. Grefenstette, ed., Morgan-Kaufmann, San Francisco, 1987, pp. 82–89.
[111] J. H. Holland, *Adaptation in Natural and Artificial Systems*, MIT Press, Cambridge, MA, 1993.
[112] W. Hordijk, *A measure of landscapes*, Evol. Comp., 4 (1996), pp. 335–360.
[113] W. Hordijk, *Correlation analysis of the synchronizing-CA landscape*, Phys. D, 107 (1997), pp. 255–264.
[114] W. Hordijk and P. F. Stadler, *Amplitude spectra of fitness landscapes*, Adv. Complex Systems, 1 (1998), pp. 39–66.
[115] H. Hotelling, *Analysis of a complex of statistical variables into principal components*, J. Educ. Psych., 24 (1933), pp. 417–441 and 498–520.
[116] J. A. Howell, T. F. Smith, and M. S. Waterman, *Computation of generating functions for biological molecules*, SIAM J. Appl. Math., 39 (1980), pp. 119–133.
[117] M. A. Huynen, *Exploring phenotype space through neutral evolution*, J. Mol. Evol., 43 (1996), pp. 165–169.

[118] M. A. Huynen, P. F. Stadler, and W. Fontana, *Smoothness within ruggedness: The role of neutrality in adaptation*, Proc. Natl. Acad. Sci. USA, 93 (1996), pp. 397–401.
[119] T. Jones, *One Operator, One Landscape*, Tech. Report #95-02-025, Santa Fe Institute, Santa Fe, NM, 1995.
[120] T. Jones and S. Forrest, *Fitness distance correlation as a measure of problem difficulty for genetic algorithms*, in Proceedings of the Sixth International Conference on Genetic Algorithms, L. Eshelman, ed., Morgan-Kaufmann, San Mateo, CA, 1995, pp. 184–192.
[121] G. F. Joyce, *Directed molecular evolution*, Sci. Amer., 267 (1992), pp. 48–55.
[122] K. Karhunen, *Zur Spektraltheorie Stochasticher Prozesse*, Ann. Acad. Sci. Fenn. Ser. A. I. Math.-Phys., 34 (1947), pp. 1–7.
[123] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, Academic Press, New York, 1975.
[124] S. A. Kauffman, *The Origin of Order*, Oxford University Press, New York, Oxford, 1993.
[125] S. A. Kauffman and S. Levin, *Towards a general theory of adaptive walks on rugged landscapes*, J. Theoret. Biol., 128 (1987), pp. 11–45.
[126] D. C. Kent, *Convergence functions and their related topologies*, Fund. Math., 54 (1964), pp. 125–133.
[127] D. C. Kent, *A note on pretopologies*, Fund. Math., 62 (1968), pp. 95–100.
[128] W. Kern, *On the depth of combinatorial optimization problems*, Discrete Appl. Math., 43 (1993), pp. 115–129.
[129] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Optimization by simulated annealing*, Science, 220 (1983), pp. 671–680.
[130] D. J. Klein and D. Babić, *Partial orderings in chemistry*, J. Chem. Inf. Comput. Sci., 37 (1997), pp. 656–671.
[131] M. Klin, C. Rücker, G. Rücker, and G. Tinhofer, *Algebraic combinatorics in mathematical chemistry. Methods and algorithms.* I. *Permutation groups and coherent (cellular) algrebras*, MATCH, 40 (1999), pp. 7–138.
[132] M. H. Klin, R. Pöschel, and K. Rosenbaum, *Angewandte Algebra*, Vieweg, Braunschweig, 1988 (in German).
[133] S. J. Klug and M. Famulok, *All you wanted to know about SELEX*, Mol. Biol. Rep., 20 (1994), pp. 97–107.
[134] K. Kollman, J. H. Miller, and S. E. Page, *Adaptive parties in spatial elections*, Amer. Pol. Sci. Rev., 86 (1992), pp. 929–937.
[135] S. Kopp and C. M. Reidys, *Neutral networks: A combinatorial perspective*, Adv. Complex Systems, 2 (1999), pp. 283–301.
[136] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press, Cambridge, MA, 1992.
[137] H. Kubinyi, *QSAR: Hansch Analysis and Related Approaches*, Methods and Principles in Medicinal Chemistry, Vol. 1, VCH, Weinheim, 1995.
[138] P. Larrañaga, C. M. H. Kuijpers, R. H. Murga, I. Inza, and S. Dizdarevic, *Genetic algorithms for the travelling salesman problem: A review of representations and operators*, Artificial Intelligence Rev., 13 (1999), pp. 129–170.
[139] E. L. Lawler, J. K. Lenstra, A. H. G. R. Kan, and D. B. Shmoys, *The Traveling Salesman Problem. A Guided Tour of Combinatorial Optimization*, John Wiley, New York, 1985.
[140] D. A. Levinthal, *Adaptation on rugged landscapes*, Manage. Sci., 43 (1997), pp. 934–950.
[141] G. E. Liepins and M. D. Vose, *Deceptiveness and genetic algorithm dynamics*, in Foundations of Genetic Algorithms, G. J. E. Rawlins, ed., Morgan-Kaufmann, San Mateo, CA, 1991, pp. 36–50.
[142] G. E. Liepins and M. D. Vose, *Polynomials, basis sets, and deceptiveness in genetic algorithms*, Complex Systems, 5 (1991), pp. 45–61.
[143] D. L. Miller and J. F. Pekny, *Exact solution of large asymmetric traveling salesman problems*, Science, 251 (1991), pp. 754–761.
[144] M. M. Loève, *Probability Theory*, VanNostrand, Princeton, NJ, 1955.
[145] L. Lovasz, *Random walks on graphs: A survey*, in Combinatorics, Paul Erdős is Eighty, Vol. 2, Janos Bolyai Mathematical Society, Budapest, 1996, pp. 353–398.
[146] C. A. Macken, P. S. Hagan, and A. S. Perelson, *Evolutionary walks on rugged landscapes*, SIAM J. Appl. Math., 51 (1991), pp. 799–827.
[147] C. A. Macken and A. S. Perelson, *Protein evolution on rugged landscapes*, Proc. Natl. Acad. Sci. USA, 86 (1989), pp. 6191–6195.
[148] K. Malarz and D. Tiggemann, *Dynamics in Eigen quasispecies model*, Internat. J. Modern Phys. C, 9 (1998), pp. 481–490.

[149] D. Maslen and D. Rockmore, *Generalized FFTs—a survey of some recent results*, in Groups and Computation II, L. Finkelstein and W. Kantor, eds., DIMACS Ser. Discrete Math Theoret. Comput. Sci. 28, AMS, Providence, RI, 1996, pp. 183–238.

[150] R. Merris, *A survey of graph Laplacians*, Linear Algebra Appl., 39 (1995), pp. 19–31.

[151] P. Merz and B. Freisleben, *Fitness landscape analysis and memetic algorithms for the quadratic assignment problem*, IEEE Trans. Evol. Comput., 4 (2000), pp. 337–352.

[152] M. Mézard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond*, World Scientific, Singapore, 1987.

[153] P. G. Mezey, *Potential Energy Hypersurfaces*, Elsevier, Amsterdam, 1987.

[154] D. R. Mills, R. L. Peterson, and S. Spiegelman, *An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule*, Proc. Natl. Acad. Sci. USA, 58 (1967), pp. 217–224.

[155] J. Milne, *Étale Cohomology*, Princeton University Press, Princeton, NJ, 1980.

[156] M. Mitchell, *An Introduction to Genetic Algorithms*, 2nd ed., MIT Press, Cambridge, MA, 1998.

[157] B. Mohar, *The Laplacian spectrum of graphs*, in Graph Theory, Combinatorics, and Applications, Y. Alavi, G. Chartrand, O. Ollermann, and A. Schwenk, eds., John Wiley, New York, 1991, pp. 871–898.

[158] B. Mohar, *Some applications of Laplace eigenvalues of graphs*, in Graph Symmetry: Algebraic Methods and Applications, G. Hahn and G. Sabidussi, eds., NATO Adv. Sci. Inst. Ser. C 497, Kluwer Academic, Dordrecht, The Netherlands, 1997, pp. 227–275.

[159] P. A. P. Moran, *Random processes in genetics*, Proc. Camb. Phil. Soc., 54 (1958), pp. 60–71.

[160] S. R. Morgan and P. G. Higgs, *Barrier heights between groundstates in a model of RNA secondary structure*, J. Phys. A, 31 (1998), pp. 3153–3170.

[161] H. Mortveit and C. M. Reidys, *Discrete, sequential dynamical systems*, Discrete Math., 226 (2001), pp. 281–295.

[162] B. Naudts and L. Kallel, *A comparison of predictive measures of problem difficulty in evolutionary algorithms*, IEEE Trans. Evol. Comp., 4 (2000), pp. 1–15.

[163] K. Nemoto, *Metastable states of the SK spin glass model*, J. Phys. A, 21 (1988), pp. L287–L294.

[164] A. Neumaier, *Molecular modeling of proteins and mathematical prediction of protein structure*, SIAM Rev., 39 (1997), pp. 407–460.

[165] M. E. J. Newman and R. Engelhardt, *Effects of selective neutrality on the evolution of molecular species*, Proc. Roy. Soc. London Ser. B, 265 (1998), pp. 1333–1338.

[166] K. J. Niklas, *Adaptive walks through fitness landscapes for early vascular land plants*, Amer. J. Botany, 84 (1997), pp. 16–25.

[167] E. A. Nix and M. D. Vose, *A Markov chain analysis on a simple genetic algorithm*, IEEE Trans. Systems, Man, Cybern., SMC-25 (1992), pp. 655–659.

[168] J. R. Norris, *Markov Chains*, Cambridge University Press, Cambridge, UK, 1997.

[169] M. Nowak and P. Schuster, *Error thresholds of replication in finite populations, mutation frequencies and the onset of Muller's ratchet*, J. Theoret. Biology, 137 (1989), pp. 375–395.

[170] J. N. Onuchic, H. Nymeyer, A. E. Garcia, J. Chahine, and N. D. Socci, *The energy landscape theory of protein folding: Insights into folding mechanisms and scenarios*, Adv. Protein Chem., 53 (2000), pp. 87–152.

[171] R. Otten and L. van Ginneken, *The Annealing Algorithm*, Kluwer Academic, Boston, 1989.

[172] S. E. Page and D. E. Richardson, *Walsh functions, scheme variance, and deception*, Complex Systems, 6 (1992), pp. 125–136.

[173] R. Palmer, *Optimization on rugged landscapes*, in Molecular Evolution on Rugged Landscapes: Proteins, RNA, and the Immune System, A. S. Perelson and S. A. Kauffman, eds., Addison–Wesley, Redwood City, CA, 1991, pp. 3–25.

[174] A. S. Perelson and S. A. Kauffman, eds., *Molecular Evolution on Rugged Landscapes: Proteins, RNA, and the Immune System*, Santa Fe Inst. Stud. 9, Addison–Wesley, Reading, MA, 1991.

[175] A. S. Perelson and C. A. Macken, *Protein evolution on partially correlated landscapes*, Proc. Natl. Acad. Sci. USA, 92 (1995), pp. 9657–9661.

[176] D. L. Powers, *Graph partitioning by eigenvectors*, Linear Algebra Appl., 101 (1988), pp. 121–133.

[177] R. Rammal, G. Toulouse, and M. A. Virasoro, *Ultrametricity for physicists*, Rev. Modern Phys., 58 (1986), pp. 765–788.

[178] I. Rechenberg, *Evolutionstrategie*, Frommann-Holzboog, Stuttgart, 1973.
[179] C. M. Reidys, *Random induced subgraphs of generalized n-cubes*, Adv. in Appl. Math., 19 (1997), pp. 360–377.
[180] C. M. Reidys, *Acyclic orientations of random graphs*, Adv. in Appl. Math., 21 (1998), pp. 181–192.
[181] C. M. Reidys, *Cohomology of Landscapes*, manuscript.
[182] C. M. Reidys, *Random graphs and sequence to structure maps*, Combin. Probab. Comput., submitted.
[183] C. M. Reidys, *Random structures*, Ann. Comb., 4 (2000), pp. 375–382.
[184] C. M. Reidys, *SDS Phase Space Properties*, manuscript.
[185] C. M. Reidys, *Acyclic orientations and sequential dynamical systems*, Adv. in Appl. Math., to appear.
[186] C. M. Reidys, C. V. Forst, and P. Schuster, *Replication and mutation on neutral networks of RNA secondary structures*, Bull. Math. Biol., 63 (2001), pp. 57–94.
[187] C. M. Reidys and P. F. Stadler, *Bio-molecular shapes and algebraic structures*, Computers and Chemistry, 20 (1996), pp. 85–94.
[188] C. M. Reidys and P. F. Stadler, *Neutrality in fitness landscapes*, Appl. Math. Comput., 117 (2001), pp. 321–350.
[189] C. M. Reidys, P. F. Stadler, and P. Schuster, *Generic properties of combinatory maps: Neural networks of RNA secondary structures*, Bull. Math. Biol., 59 (1997), pp. 339–397.
[190] H. Rieger, *The number of solutions of the Thouless-Anderson-Palmer equations for p-spin interaction spin glasses*, Phys. Rev. B, 46 (1992), pp. 14655–14661.
[191] D. Rockmore, *Some applications of generalized FFTs*, in Groups and Computation II, L. Finkelstein and W. Kantor, eds., DIMACS Ser. Discrete Math. Theoret. Comput. Sci. 28, AMS, Providence, RI, 1995, pp. 329–370.
[192] D. Rockmore, P. Kostelec, W. Hordijk, and P. F. Stadler, *Fast Fourier transform for fitness landscapes*, Appl. Comput. Harmon. Anal., to appear.
[193] G. Rota, *On the foundations of combinatorial theory* I: *Theory of Möbius functions*, Z. Wahrscheinl. u. verw. Gebiete, 2 (1964), pp. 340–368.
[194] G. Rudolph, *Finite Markov chain results in evolutionary computation: A tour d'horizon*, Fund. Inform., 35 (1998), pp. 67–89.
[195] J. Ryan, *The depth and width of local minima in discrete solution spaces*, Discrete Appl. Math., 56 (1995), pp. 75–82.
[196] M. Sassanfar and J. W. Szostak, *An RNA motif that binds ATP*, Nature, 364 (1993), pp. 550–553.
[197] W. R. Schmitt and M. S. Waterman, *Plane trees and RNA secondary structure*, Discrete Appl. Math., 51 (1994), pp. 317–323.
[198] E. A. Schultes and D. P. Bartes, *One sequence, two ribozymes: Implications for the emergence of new ribozyme folds*, Science, 289 (2000), pp. 448–452.
[199] P. Schuster, *Landscapes and molecular evolution*, Phys. D, 107 (1997), pp. 351–365.
[200] P. Schuster, W. Fontana, P. F. Stadler, and I. L. Hofacker, *From sequences to shapes and back: A case study in RNA secondary structures*, Proc. Roy. Soc. London Ser. B, 255 (1994), pp. 279–284.
[201] P. Schuster and P. F. Stadler, *Landscapes: Complex optimization problems and biopolymer structures*, Computers and Chemistry, 18 (1994), pp. 295–314.
[202] P. Schuster and J. Swetina, *Stationary mutant distributions and evolutionary optimization*, Bull. Math. Biol., 50 (1988), pp. 635–660.
[203] M. Shpak and G. P. Wagner, *Asymmetry of configuration spaces induced by unequal crossover: Implications for a mathematical theory of evolutionary innovation*, Artificial Life, 6 (2000), pp. 25–43.
[204] G. B. Sorkin, *Combinatorial Optimization, Simulated Annealing, and Fractals*, Tech. Report RC13674 (No. 61253), IBM Research Report, White Plains, NY, 1988.
[205] W. M. Spears, *Evolutionary Algorithms: The Role of Mutation and Recombination*, Springer-Verlag, Heidelberg, 2000.
[206] B. M. Stadler, *Adaptive platform dynamics in multi-party spatial voting*, Adv. Complex Systems, 2 (1999), pp. 101–116.
[207] B. M. R. Stadler and P. F. Stadler, *Generalized topological spaces in evolutionary theory and combinatorial chemistry*, J. Chem. Inf. Comput. Sci., to appear.
[208] B. M. R. Stadler, P. F. Stadler, W. Fontana, and G. P. Wagner, *The topology of the possible: Formal spaces underlying patterns of evolutionary change*, J. Theor. Biol., 213 (2001), pp. 241–274.

[209] P. F. Stadler, *Random walks and orthogonal functions associated with highly symmetric graphs*, Discrete Math., 145 (1995), pp. 229–238.
[210] P. F. Stadler, *Towards a theory of landscapes*, in Complex Systems and Binary Networks, R. Lopéz-Peña, R. Capovilla, R. García-Pelayo, H. Waelbroeck, and F. Zertuche, eds., Springer-Verlag, Berlin, New York, 1995, pp. 77–163.
[211] P. F. Stadler, *Landscapes and their correlation functions*, J. Math. Chem., 20 (1996), pp. 1–45.
[212] P. F. Stadler, *Fitness landscapes arising from the sequence-structure maps of biopolymers*, J. Mol. Struct. (THEOCHEM), 463 (1999), pp. 7–19.
[213] P. F. Stadler, *Fitness landscapes*, in Biological Evolution and Statistical Physics, M. Lässing and A. Valleriani, eds., Springer-Verlag, Heidelberg, to appear.
[214] P. F. Stadler, *Spectral landscape theory*, in Evolutionary Dynamics—Exploring the Interplay of Selection, Neutrality, Accident, and Function, J. P. Crutchfield and P. Schuster, eds., Oxford University Press, New York, to appear.
[215] P. F. Stadler and R. Happel, *Correlation structure of the landscape of the graph-bipartitioning-problem*, J. Phys. A Math. Gen., 25 (1992), pp. 3103–3110.
[216] P. F. Stadler, *Random field models for fitness landscapes*, J. Math. Biol., 38 (1999), pp. 435–478.
[217] P. F. Stadler and B. Krakhofer, *Local minima of p-spin models*, Rev. Mexicana Fís., 42 (1996), pp. 355–363.
[218] P. F. Stadler and W. Schnabl, *The landscape of the travelling salesman problem*, Phys. Lett. A, 161 (1992), pp. 337–344.
[219] P. F. Stadler, R. Seitz, and G. P. Wagner, *Population dependent Fourier decomposition of fitness landscapes over recombination spaces: Evolvability of complex characters*, Bull. Math. Biol., 62 (2000), pp. 399–428.
[220] P. F. Stadler and G. Tinhofer, *Equitable partitions, coherent algebras and random walks: Applications to the correlation structure of landscapes*, MATCH, 40 (1999), pp. 215–261.
[221] P. F. Stadler and G. P. Wagner, *The algebraic theory of recombination spaces*, Evol. Comp., 5 (1998), pp. 241–275.
[222] R. P. Stanley, *Acyclic orientations of graphs*, Discrete Math., 5 (1973), pp. 171–178.
[223] C. R. Stephens and H. Waelbroeck, *Effective degrees of freedom in genetic algorithms*, Phys. Rev. E, 57 (1998), pp. 3251–3264.
[224] J. Swetina and P. Schuster, *Self-replication with errors—a model for polynucleotide replication*, Biophys. Chem., 16 (1982), pp. 329–345.
[225] M. Tacker, P. F. Stadler, E. G. Bornberg-Bauer, I. L. Hofacker, and P. Schuster, *Algorithm independent properties of RNA structure prediction*, Eur. Biophys. J., 25 (1996), pp. 115–130.
[226] F. Tanaka and S. F. Edwards, *Analytic theory of the ground state properties of a spin glass: I. Ising spin glass*, J. Phys. F Metal Phys., 10 (1980), pp. 2769–2778.
[227] P. Tarazona, *Error-thresholds for molecular quasi-species as phase transitions: From simple landscapes to spinglass models*, Phys. Rev. A, 45 (1992), pp. 6038–6050.
[228] D. J. Thouless, P. W. Anderson, and R. G. Palmer, *Solution of "Solvable model of a spin glass,"* Phil. Mag., 35 (1977), pp. 593–601.
[229] C. Tuerk and L. Gold, *Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase*, Science, 249 (1990), p. 505.
[230] H. van der Holst, *Topological and Spectral Graph Characterizations*, Ph.D. thesis, Universiteit van Amsterdam, 1996.
[231] E. Van Nimwegen and J. P. Crutchfield, *Metastable evolutionary dynamics: Crossing fitness barriers or escaping via neutral paths?*, Bull. Math. Biol., 62 (2000), pp. 799–848.
[232] V. K. Vassilev, J. F. Miller, and T. C. Fogarty, *On the nature of two-bit multiplier landscapes*, in Proceedings of the 1st NASA/DoD Workshop on Evolvable Hardware, A. Stoica, D. Keymeulen, and J. Lohn, eds., IEEE Computer Society, Los Alamitos, CA, 1999, pp. 36–45.
[233] A. M. Vertechi and M. A. Virasoro, *Energy barriers in SK spin glass models*, J. Phys. France, 50 (1989), pp. 2325–2332.
[234] M. D. Vose and A. H. Wright, *The simple genetic algorithm and the Walsh transform. Part* I: *Theory*, Evol. Comp., 6 (1998), pp. 253–274.
[235] M. D. Vose and A. H. Wright, *The simple genetic algorithm and the Walsh transform. Part* II: *The inverse*, Evol. Comp., 6 (1998), pp. 275–289.
[236] G. P. Wagner and P. Krall, *What is the difference between models of error thresholds and Muller's ratchet?*, J. Math. Biol., 32 (1993), pp. 33–44.

[237] G. P. Wagner and P. F. Stadler, *Complex adaptations and the structure of recombination spaces*, in Algebraic Engineering, C. Nehaniv and M. Itô, eds., World Scientific, Singapore, 1999, pp. 96–115.
[238] D. J. Wales, J. P. K. Doye, M. A. Miller, P. N. Mortenson, and T. P. Walsh, *Energy landscapes: From clusters to biomolecules*, Adv. Chem. Phys., 115 (2000), pp. 3–111.
[239] D. J. Wales, M. A. Miller, and T. R. Walsh, *Archetypal energy landscapes*, Nature, 394 (1998), pp. 758–760.
[240] J. Wang and M. Xu, *Quasi-abelian Cayley graphs and Parsons graphs*, European J. Combin., 18 (1997), pp. 597–600.
[241] M. S. Waterman, *Combinatorics of RNA hairpins and cloverleaves*, Stud. Appl. Math., 60 (1978), pp. 91–96.
[242] E. D. Weinberger, *Correlated and uncorrelated fitness landscapes and how to tell the difference*, Biol. Cybern., 63 (1990), pp. 325–336.
[243] E. D. Weinberger, *Fourier and Taylor series on fitness landscapes*, Biol. Cybern., 65 (1991), pp. 321–330.
[244] E. D. Weinberger, *Local properties of Kauffman's N-k model: A tunably rugged energy landscape*, Phys. Rev. A, 44 (1991), pp. 6399–6413.
[245] B. Y. Weisfeiler and A. A. Leman, *Reduction of a graph to a canonical form and an algebra arising during this reduction*, Naucho–Techn. Inf. Ser. 2, 9 (1968), pp. 12–16 (in Russian).
[246] F. H. Westheimer, *Polyribonucleic acids as enzymes*, Nature, 319 (1986), pp. 534–536.
[247] L. D. Whitley, *Fundamental principles of deception in genetic search*, in Foundations of Genetic Algorithms, G. Rawlins, ed., Morgan-Kaufmann, San Mateo, CA, 1991, pp. 221–241.
[248] T. Wiehe, *Model dependency of error thresholds: The role of fitness functions and contrasts between the finite and infinite sites models*, Genetical Research, 69 (1997), pp. 127–136.
[249] C. O. Wilke, *Adaptive evolution on neutral networks*, Bull. Math. Biol., 63 (2001), pp. 715–730.
[250] G. Woodcock and P. G. Higgs, *Population evolution on a multiplicative single-peak fitness landscape*, J. Theoret. Biology, 179 (1996), pp. 61–73.
[251] S. Wright, *The roles of mutation, inbreeding, crossbreeding and selection in evolution*, in Proceedings of the Sixth International Congress on Genetics, Vol. 1, D. F. Jones, ed., Brooklyn Botanic Garden, Menasha, WI, 1932, pp. 356–366.
[252] S. Wright, *"Surfaces" of selective value*, Proc. Natl. Acad. Sci. USA, 58 (1967), pp. 165–172.
[253] B. Zgrablic, *On adjacency-transitive graphs*, Discrete Math., 182 (1998), pp. 321–332.